Summer 7-12-2011

Remarkable Features of the Hotdog-fold Thioesterases Involved in Phylloquinone (vitamin K1) Biosynthesis

Joshua R. Widhalm
University of Nebraska-Lincoln

REMARKABLE FEATURES OF THE HOTDOG-FOLD THIOESTERASES INVOLVED IN PHYLLOQUINONE (VITAMIN K_1) BIOSYNTHESIS

by

Joshua R. Widhalm

A DISSERTATION

Presented to the Faculty of

The Graduate College at the University of Nebraska

In Partial Fulfillment of Requirements

For the Degree of Doctor of Philosophy

Major: Agronomy

Under the Supervision of Professor Gilles J.C. Basset

Lincoln, Nebraska

August, 2011

REMARKABLE FEATURES OF THE HOTDOG-FOLD THIOESTERASES INVOLVED IN PHYLLOQUINONE (VITAMIN K_1) BIOSYNTHESIS

Joshua Rodney Widhalm, Ph.D.

University of Nebraska, 2011

Advisor: Gilles J.C. Basset

Phylloquinone (vitamin K_1) is a bipartite molecule, consisting of a naphthoquinone ring attached to a phytyl side chain, that is synthesized by plants and certain cyanobacteria to serve as an electron carrier in photosystem I. The coupling of the ring and isoprenyl moieties relies on the cleavage of the CoA-thioester linkage with 1,4-dihydroxy-2-naphthoate (DHNA). It has long been a mystery if this hydrolysis is an enzymatic or chemical process. Using comparative genomics, protein biochemistry, genetics and metabolic profiling, we identified a cyanobacterial thioesterase responsible for the *in vivo* hydrolysis of DHNA-CoA. This enzyme bears a signature domain of the 4-hydroxybenzoyl-CoA thioesterase (4HBT) family of Hotdog-fold proteins.

Surprisingly, plants, which obtained most of their phylloquinone biosynthetic genes with the acquisition of the plastid, do not contain orthologs of cyanobacterial DHNA-CoA thioesterase. We tested all of the predicted 4HBT Hotdog-fold proteins in Arabidopsis by functional complementation of the cyanobacterial mutant. We found two genes encoding functional DHNA-CoA thioesterases that display low percentages of identity and dissimilar catalytic motifs from their cyanobacterial counterparts. It appears that plant DHNA-CoA thioesterases originated from a horizontal gene transfer with a species of the *Lactobacillales* order. The cognate T-DNA knockout lines exhibit reduced DHNA-CoA thioesterase activity and phylloquinone content. Fluorescently tagging the *Arabidopsis* enzymes revealed that they are localized to the peroxisome. Subcellular fractionation

assays confirmed this providing the first biochemical evidence for the involvement of peroxisomes in phylloquinone biosynthesis.

Recent proteomics and GFP-reporter projects suggest that the two steps preceding DHNA-CoA thioesterase are also peroxisomal. Thus, the current model of phylloquinone biosynthesis reflects a split between plastids and peroxisomes, implying the movement of intermediates between the organelles. To assess the importance of the cognate transport steps, we have re-routed the peroxisomal branch of the pathway to plastids in *Camelina sativa*. Here we report the findings of our metabolic engineering strategy on the pool of phylloquinone.

ACKNOWLEDGMENTS

First of all, I express sincere gratitude to my advisor Dr. Gilles Basset for his guidance over the past four years. He takes his responsibility to rigorously prepare the next generation of scientists very seriously. He has taught me many important concepts and techniques, but he has also made me understand the value of high standards, accountability, and taking pride in my work. Should I go on to have a successful career, it will be rooted in the mentoring I received while working in his lab. I also want to acknowledge the role the Post-Docs in his lab, Chloë, Fabienne, and Anne-Lise, have played in training me. I have learned a lot from each of them, and appreciate their enthusiasm and support for my work.

I am also grateful to the members of my supervisory committee, they have been instrumental to my education. I have had the great pleasure not only to take courses from most of them, but also have direct interactions outside of the classroom. I have been at Nebraska now for ten years, and my time with Dr. John Markwell goes back almost to the beginning: I thank him for seeing me through from the first time I used a pipette to the end of my Ph.D. I am indebted to Dr. Sally Mackenzie for bringing me to the Center for Plant Science Innovation and connecting me with Dr. Basset's lab. I am also thankful to her and Dr. Don Weeks for serving as readers for my dissertation. In addition, I appreciate all of the support and advice I have received from Dr. Ed Cahoon and his lab on many projects the last several years.

I am thankful to the Department of Agronomy and Horticulture for funding for my research assistantship. I would also like to thank the graduate students, Post-Docs, and staff in the PSI and the Biochemistry and Agronomy and Horticulture Departments for their engaging scientific conversations and technical assistance.

Last, and certainly not least, I am thankful for the love and support I have received from my family. I am especially grateful to my wife Melissa. She has been by my side throughout, always encouraging and believing in me. Thanks also to my little boy Reid. His contagious jovial spirit helps me to forget the stresses of the day and start anew the next.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

Literature Review

Preface

In the early half of the 20th century it was observed that chicks fed a reconstituted "sterol-free" diet developed a hemorrhagic disease as a result of the inability to coagulate blood (1-7). The absence of fats, sterols, and then-known vitamins, was dismissed as the underlying cause of the disease after subsequent feeding experiments revealed that reintroduction of these compounds failed to stop the disease. The supplementation of the chicks' diet with various plant or animal products such as fresh cabbage, dried alfalfa, tomatoes, hemp seeds, and putrefied fish meal (but not fresh) was able to stop the hemorrhaging (4, 5, 7). Soon thereafter it was discovered that the green parts of plants, as well as the microorganisms in the putrefied fish, were the source of the antihemorrhagic factor (5). This factor was coined "vitamin K" in 1935, because it was the next available letter in the alphabet for the systematic naming of vitamins (7). Serendipitously, "K" happens to also be the first letter in the word "koagulation", which is the Danish spelling of "coagaulation", and the native language of Henrik Dam, one of the co-discoverers of vitamin K. By the early 1940's, the structures of the various forms of vitamin K were elucidated and able to be chemically synthesized (8-12). It took another 30 years to determine the precursors needed to synthesize vitamin K, and it was not until this last decade that many of the enzymes involved in its biosynthesis were identified in plants.

In the strictest sense, the usage of the term *vitamin K* should be reserved for describing the vital cofactor obtained through the diet by vertebrates. Throughout my dissertation, however, I will sometimes use *vitamin K* as shorthand to collectively refer to phylloquinone and the menaquinones. It is also worth noting that in the literature, particularly in medical journals, it is common to see the use of *vitamin* K_1 in place of phylloquinone and *vitamin* K_2 for the various menaquinones.

In this chapter I will summarize the literature surrounding the function and biosynthesis of vitamin K. Many of the figures and the spirit of this chapter come from a recent review we wrote for the *Advances of Botanical Research*:

Phylloquinone (Vitamin K_1): Function, Enzymes and Genes

Chloë van Oostende, Joshua R. Widhalm, Fabienne Furt, Anne-Lise Ducluzeau and Gilles J. Basset

An invited review for the *Advances in Botanical Research* that will be published in 2011

Chemistry of Vitamin K

"Vitamin K" is the term used for the class of fat-soluble compounds that consist of a naphthoquinone ring attached to a variable poly-isoprenyl side chain. The principle forms of vitamin K are vitamin K_1 (phylloquinone: 2-methyl-3-phytyl-1,4-naphthoquinone) and vitamin K_2 (menaquinones: 2-methyl-3-transpolyprenyl-1,4-naphthoquinone; Fig. 1-1A). The side chain of phylloquinone is partially unsaturated, made up of one isopentenyl, followed by three isopentyl units (*i.e.* C20; Fig. 1-1A). Menaquinones, on the other hand, have fully unsaturated side chains solely comprised of *n*=2-13 isopentenyl units (*i.e.* C10-C65; Fig. 1-1A).

The naphthoquinone ring moiety of vitamin K can exist in various redox forms, ranging from the epoxide (most oxidized) to the quinol (most reduced) with quinone and semi-quinone intermediates (Fig. 1-1B). The epoxide form has only been reported so far in animals as a byproduct of certain enzymatic reactions. In some phylloquinone-producing organisms both the quinone and quinol have been detected, with the former making up the majority of the metabolite pool (13-15).

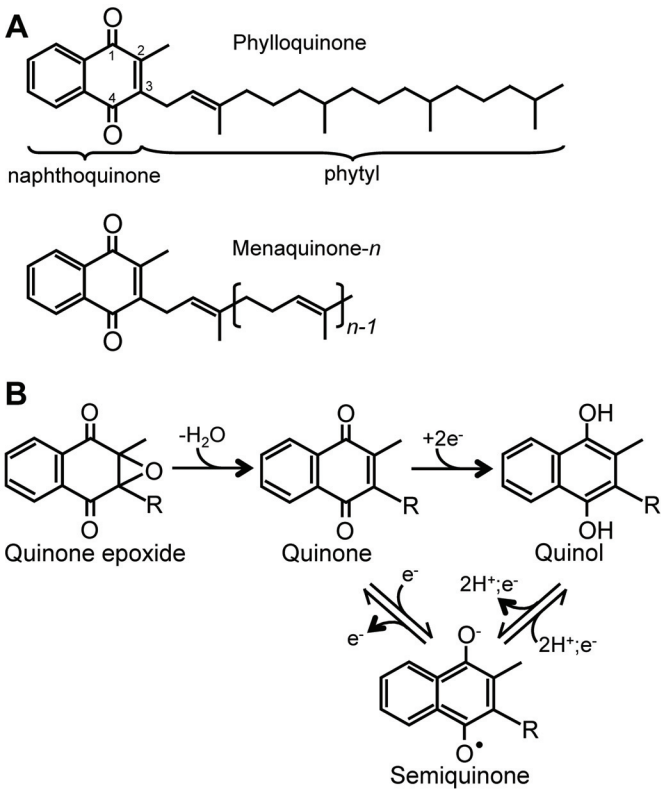


Fig. 1-1. Structures and redox forms of vitamin K. (*A*) Stuctures of phylloquinone and menaquinones. The phytyl side chain of phylloquinone contains one isopentenyl unit followed by three isopentyl units, while menaquinones have one comprised of all isopentyl units up to *n*=13. (*B*) Different inter-convertible redox forms of the naphthoquinone ring. R: poly-isoprenyl moiety.

Occurrence and Functions of Vitamin K

Vertebrates: As the name implies, vitamin K is not synthesized by vertebrates and therefore must be obtained through the diet (as phylloquinone and/or menaquinone). It is used as a cofactor in the γ-carboxylation of specific glutamate residues to confer strong chelating properties to certain proteins with activities depending on binding calcium. Examples include some of the prothrombin blood coagulation factors and osteocalcin, which participates in bone metabolism. Vitamin K must be in the quinol form to serve as a cofactor for γ-carboxylase, and after donating its electrons, is converted to the metabolically inactive epoxide form. The Vitamin K epOxide Reductase (VKOR) integral enzyme complex contains a catalytic subunit (VKORC1; EC 1.1.4.1) responsible for reducing the epoxide back to the quinone, and the quinone to the quinol form (Fig. 1-2; 16). The anticoagulant drug warfarin, often prescribed to prevent blood clots or administered

to heart attack patients (and the active ingredient in many commercial rodenticides!), is an antagonist of VKOR.

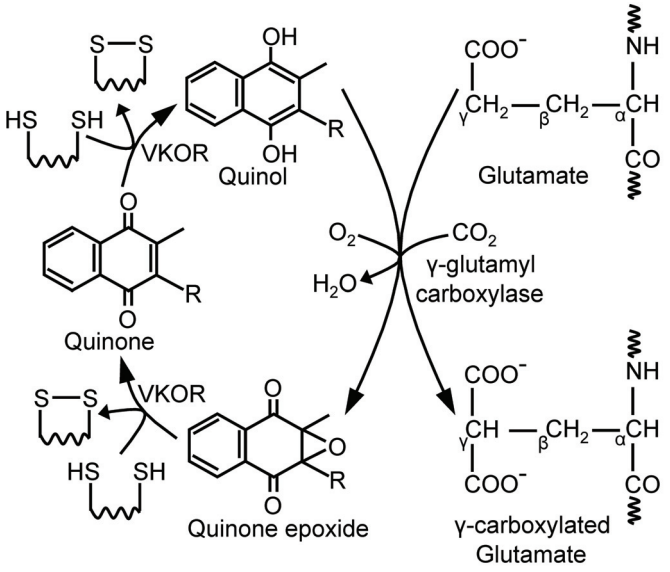


Fig. 1-2: Scheme of the vitamin K-dependent γ-carboxylation of glutamyl residues in vertebrates. Vitamin K quinol is converted to an oxygenated intermediate that abstracts a proton from the γ-carbon of the glutamyl residue of the carboxylase substrate, followed by the addition of carbon dioxide. The concomitant oxidation of vitamin K quinol into vitamin K epoxide, and its subsequent salvaging by the enzyme complex vitamin K epoxide reductase (VKOR), is called the vitamin K cycle.

(Most) Archaea and Bacteria: Menaquinones are still found in most groups of extant archaea and bacteria (17). They are the ancestral quinone used in electron transport, evolving prior to oxygenic photosynthesis (18). With the increase in atmospheric oxygen levels, however, the need for quinones with more negative midpoint redox potentials (*i.e.* plastoquinone and ubiquinone) soon became necessary for respiration and photosynthesis (reviewed in 18, 19). Nevertheless, in anaerobic environments today many archaea and bacteria have retained the ability to synthesize menaquinones to serve as carriers in respiratory electron transport chains with terminal electron acceptors besides oxygen (*e.g.* fumarate, nitrate, *etc.*; reviewed in 20). In other microbes, such as the purple γ-proteobacterium *Halorhodospira halophila*, menaquinone functions as an electron carrier at the Q_B-site of the photosynthetic reaction center (18).

Plants, diatoms, algae and cyanobacteria: Plants (13), green algae (21), and certain cyanobacteria (17) synthesize phylloquinone to serve as a single electron carrier (through the quinone–semi-quinone forms) at the A_1 site of photosystem I (22-24). Some species of cyanobacteria (*e.g. Gloeobacter violaceus* (25), red algae (26), and diatoms (27) utilize menaquinone instead of phylloquinone in photosytem I. As expected from its function in photosynthetic electron transport, phylloquinone content in plants is highest in leaves (Table 1-1). Accordingly, at the subcellular level, plastids account for most, if not all, of the phylloquinone in plant tissues (13, 28).

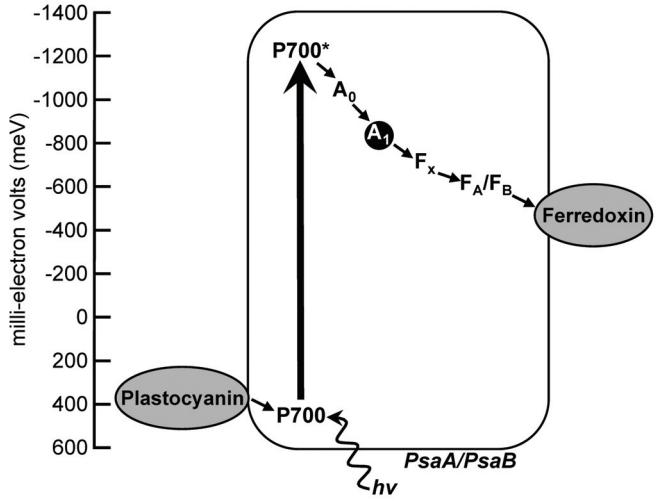


Fig. 1-3: Scheme of the electron transfer in photosystem I and approximate midpoint potentials of the cognate electron carriers in plants and cyanobacteria. Phylloquinone is located at the A_1 site of the PsaA and PsaB subunits of photosystem I to serve as a one electron carrier from chlorophyll a (A_0) to the Fe-S cluster (F_X, F_A/F_B). P700, photosystem I reaction center; P700*, excited photosystem I reaction center.

In plants, new findings suggest that half of phylloquinone is not bound to photosystem I (28, 29). Moreover, a new redox form, the quinol, has been detected in several plant species (13) and the cyanobacterium *Synechocystis* (14). Together, these suggest that phylloquinone is participating in redox reactions outside of photosystem I. One clue to an alternative function may come from the presence of mammalian-like VKORC1 proteins in plants and cyanobacteria. While there is no biochemical or genomic evidence to suggest that γ-carboxylation is occurring outside the metazoan lineage, VKORC1 homologs in plants and cyanobacteria have a C-terminal fusion with a soluble thioredoxin-like domain likely to be a protein disulfide isomerase (PDI; 15, 30-32). The fused enzyme from *Synechococcus* has been shown to couple the formation of disulfide bonds of

an artificial protein substrate *in vitro* with the reduction of phylloquinone (33). Even more, the *Arabidopsis* ortholog, encoded by *At4g35760*, is localized to plastids and can convert conjugated naphthoquinones into their quinol forms using artificial or the C-terminally fused PDI moiety as the electron donor (15). Thus, it appears that phylloquinone may be necessary for the formation of protein disulfide bonds in cyanobacteria and plant plastids.

Table 1-1: Phylloquinone content of some plant species and plant food-products. Data are compiled from (13)[(a)]; (34)[(b)]; USDA National Nutrient Database for Standard Reference (http://www.nal.usda.gov/fnic/foodcomp/search/) [(c)]. Staple crop values are shown in bold.

Species/tissue	Phylloquinone *µg/100g*
A. thaliana (green leaf)	365[(a)]
Brassica oleracea (canola oil)	127[(b)]
(collard greens)	440[(b)]
(broccoli)	180[(b)]
(brussel sprouts)	177[(b)]
(cauliflower)	20[(b)]
Cicer arietinum (chickpeas)	9[(c)]
Daucus carota (tuber)	2.7[(a)]
Lactuca sativa (green leaf)	126[(c)]
('iceberg' lettuce)	35[(b)]
Manihot esculenta (cassava)	**1.9[(c)]**
Olea europaea (olive oil)	55
Oryza sativa (grain)	**0.1[(c)]**
(green leaf)	662[(a)]
Phaseolus vulgaris (dry bean)	**5.6[(c)]**
(green beans)	33[(b)]
Solanum. lycopersicon (green leaf)	1217[(a)]
(green fruit)	19[(a)]
(red ripe fruit)	8[(a)]
Solanum tuberosum (tuber)	**1.3[(a)]**
Glycine max (soybean oil)	193[(b)]
('Edamame' seed)	31[(c)]
Triticum spp. (whole grain flour)	**1.9[(c)]**
Vicia faba (fava bean)	**9[(c)]**
Zea mays (grain)	**0.3[(c)]**
(green leaf)	1514[(a)]
(oil)	3[(b)]

Biosynthesis of Vitamin K

Background: Plants and most microorganisms use a series of nine enzymatic steps starting from chorismate to synthesize vitamin K (reactions 1-9, Fig. 1-4). This "classical pathway" is comprised of two metabolic branches, one for the naphthoquinone ring and the other for the poly-isoprenyl side chain. First, chorismate is converted to isochorismate (reaction 1, Fig. 1-4). A succinyl side chain is then added at the C2 position, followed by the β-elimination of pyruvate and aromatization of the cyclohexadiene ring (reactions 2-4, Fig. 1-4). Next, the succinyl side chain is activated by ligation with CoA and cyclized (reactions 5 and 6, Fig. 1-4). The CoA-thioester linkage is then cleaved to form 1,4-dihydroxynaphthoate (DHNA), which is conjugated to the isoprenyl moiety and methylated (reactions 7-9, Fig. 1-4). Recently an alternative pathway to synthesize menaquinone, which starts from a chorismate-inosine conjugate called futalosine, was discovered in some types of microorganisms (35), but no evidence is present for its existence in plants. This pathway will not be discussed any further.

Elucidation of vitamin K biosynthesis (the "classical pathway") began with the first radiolabeling studies showing shikimate as a precursor for menaquinone in some types of bacteria (reviewed in 36). Similar experiments suggested plants use an identical pathway with *o*-succinylbenzoate (OSB), OSB-CoA, and DHNA as intermediates (Fig. 1-4; 37-40). This work also proposed that OSB and DHNA are intermediates in the formation of anthraquinone, a building block of certain plant pigments. Additional experiments found that phytyl diphosphate and *s*-adenosylmethionine are used as substrates for prenylation and methylation of the ring, respectively (Fig. 1-4; 41, 42).

The first genetic studies on vitamin K biosynthesis began in the 1970s using forward-genetic screens in *Escherichia coli* (*E. coli*) to identify mutants unable to grow under anaerobic conditions (reviewed in 43). The affected genes were termed *menA* to *menH* (standing for men<u></u>aquinone, with the accompanying letter assigned based on the order of discovery). In Fig. 1-4 the enzymes encoded by the *men* genes are shown, but note that there is no MenG. Instead, it was shown in *E. coli* that the methyltransferase UbiE involved in ubiquinone biosynthesis doubles for menaquinone (44).

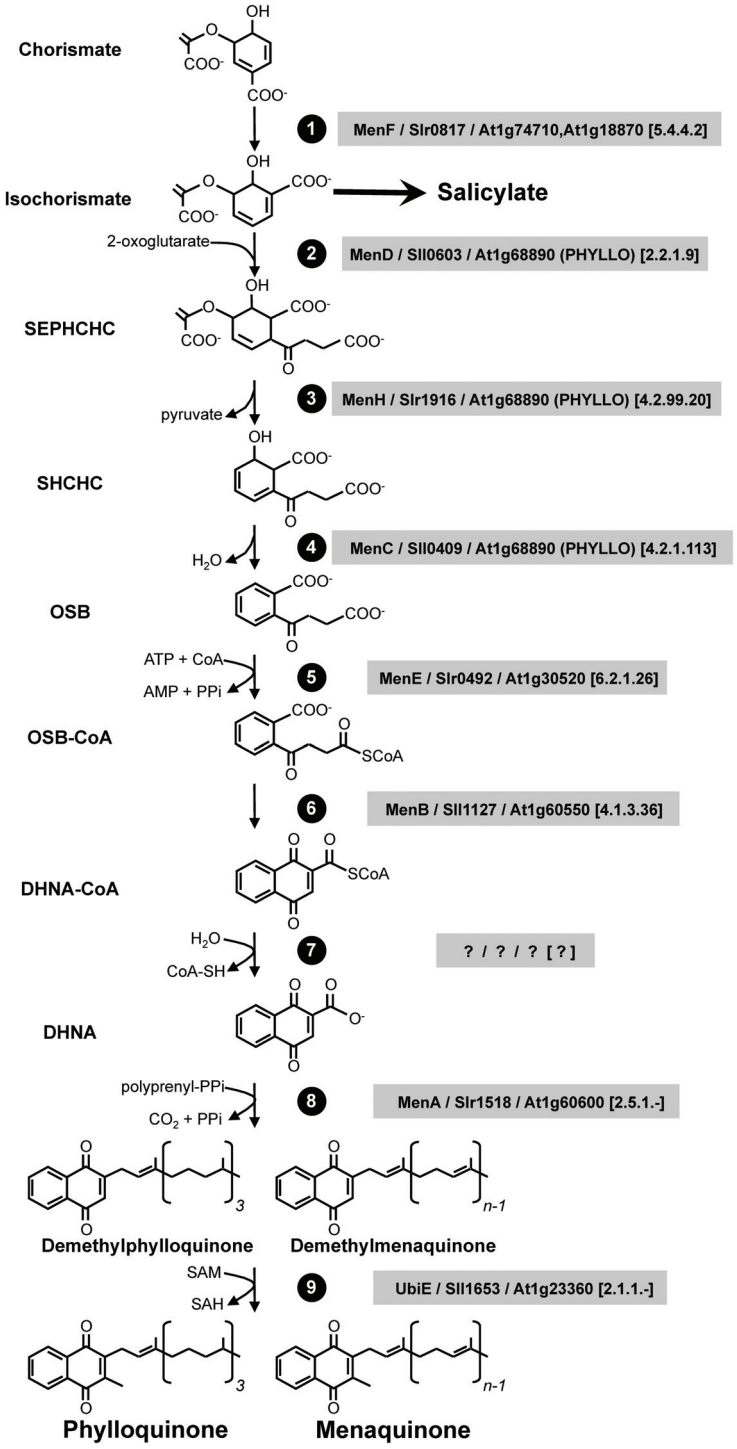


Fig. 1-4: The classically-described vitamin K biosynthetic pathway. The *E. coli* Men enzymes, and their orthologs in *Synechocystis* and *Arabidopsis*, are shown. EC numbers are indicated in brackets next to the corresponding enzymes. MenF, isochorismate synthase; MenD, SEPHCHC synthase; MenH, SHCHC synthase (note that this enzyme is currently mis-annotated in most databases as DHNA-CoA thioesterase); MenC, OSB synthase; MenE, OSB-CoA ligase; MenB, DHNA-CoA synthase; MenA, DHNA prenyltransferase; UbiE, demethylmenaquinone/demethylphylloquinone methyltransferase. SEPHCHC, 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic acid; SHCHC, (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid; SAM, *S*-adenoysl methionine; SAH, *S*-adenosyl homocysteine.

It has only been within the last ten years or so that the phylloquinone biosynthetic genes in cyanobacteria and plants have been identified. Most of these genes have been found through homology-based approaches with the *E. coli men* genes as query (see Fig. 1-4 for a list of the Men orthologs in the cyanobacterium *Synechocystis* sp. PCC 6803 and *Arabidopsis thaliana*). Nevertheless, the last two steps in formation of the naphthoquinone ring in plants remain to be characterized (reactions 6 and 7, Fig. 1-4). For the first, which is likely identical to the reaction catalyzed by MenB in *E. coli* (reaction 6, Fig. 1-4), an apparent *Arabidopsis* ortholog exists, and is just awaiting experimental confirmation. On the other hand, it is not even known if the hydrolysis of DHNA-CoA to DHNA (reaction 7, Fig. 1-4) is enzymatic. No gene encoding an enzyme with this activity has been found in any of the vitamin K-producing organisms. The rest of this section will be spent on describing each biosynthetic step in more detail. The focus will primarily be on what is known about the genes and enzymes involved in phylloquinone production in plants, and a little about the evolution of the pathway.

Isochorismate Synthase and PHYLLO (Reactions 1-4): Using a forward-genetics approach in *Arabidopsis*, it was recently discovered that plants possess a gene, named *PHYLLO* (*At1g68890*), that encodes for a single polypeptide containing four fused proteins homologous to bacterial MenF, MenD, MenC, and MenH (reactions 1-4, Fig. 1-4; 29). Interestingly, the MenF module of PHYLLO in *Arabidopsis* contains a truncated chorismate-binding domain that likely renders it unable to catalyze the formation of isochorismate (reaction 1, Fig. 1-4). This feature appears to occur throughout the monocot and dicot genomes examined so far. Instead, higher plants have evolved separate and catalytically active isochorismate synthases (reaction 1, Fig. 1-4). In *Arabidopsis*, for example, there are two isochorismate synthases (ICS1, At1g74710 and ICS2, At1g18870) that function independently of PHYLLO (45-47). The *ics1/ics2* double knockout is devoid of phylloquinone (29, 47), which is consistent with the idea that the truncated MenF domain of PHYLLO is not sufficient to produce isochorismate.

The *PHYLLO* locus is also present in the nuclear genomes of diatoms and green algae (29), though in *Chlamydomonas reinhardtii* each module may be translated as an individual

polypeptide (21). Unlike monocots and dicots, however, the MenF component of PHYLLO in these organisms contains a complete chorismate-binding domain that is *a priori* functional. The PHYLLO fusion event likely indicates that there is channeling of intermediates to synthesize OSB (Fig. 1-4). The disassociation of the isochorismate synthase activity from PHYLLO in higher plants, though, may have opened up a metabolic branch point for the plastid-synthesized hormone salicylate, which relies on the same isochorismate pool (Fig. 1-4; 45). Evidence for this comes from the *Arabidopsis ics* mutants that not only lack phylloquinone, but also salicylate (29, 47). Moreover, it was demonstrated that constitutively expressing bacterial isochorismate pyruvate-lyase, which converts isochorismate into salicylate, in tobacco chloroplasts results in increased salicylate content at the expense of phylloquinone (48).

When PHYLLO was first described it was still thought that the predicted α/β-hydrolase activity coming from the MenH domain was responsible for the hydrolysis of the DHNA-CoA thioester linkage (reaction 7, Fig. 1-4). It was also assumed then that MenD used isochorismate as substrate to directly produce 2-succinyl-6-hydroxy-2,4-cyclohexadiene (SHCHC, Fig. 1-4). Such assignments suggested that the fused enzymes within PHYLLO catalyzed nonconsecutive reactions. Shortly thereafter, however, it was shown in *E. coli* that MenD is only responsible for the addition of the succinyl side chain to isochorismate (reaction 2, Fig. 1-4; 49), and that MenH subsequently catalyzes the elimination of pyruvate from the unsuspected MenD product (reaction 3, Fig. 1-4; 50). These assignments not only demonstrated that the PHYLLO domains catalyze sequential reactions in the pathway, but they also reopened the search for how DHNA-CoA is converted to DHNA.

The CoA reactions (Reactions 5-7): After formation of OSB by PHYLLO, the next three steps in vitamin K biosynthesis complete the formation of the naphthoquinone ring (reactions 5-7, Fig. 1-4). First, the succinyl side chain of OSB is activated through formation of a high-energy bond with the pantetheine moiety of CoA (reaction 5, Fig. 1-4; 51). The responsible enzyme, OSB-CoA ligase, was initially identified in plants as part of a large-scale effort to characterize the CoA ligase family in *Arabidopsis* (52). It was found that the putative CoA-ligase gene *AAE14* (*Acyl-Activating*

Enzyme 14; At1g30520) co-expresses with many of the known phylloquinone biosynthetic genes (52). Subsequent analysis of T-DNA insertion lines corresponding to *At1g30520* revealed that the mutants lacked phylloquinone (52). It was also found that these mutants accumulated the assumed substrate OSB, and could be partially rescued by exogenous application of DHNA (52). Functional complementation of the *E.* coli *menE* mutant with the *At1g30520* cDNA was able to restore menaquinone production, confirming AAE14 is an OSB-CoA ligase (52).

Activation of the succinyl side chain allows OSB-CoA to be cyclized by DHNA-CoA synthase (reaction 6, Fig. 1-4). This enzyme is often misnamed “DHNA synthase” in the literature, because it was once believed that the enzyme produces DHNA. It has now been well established, though, that the product of the reaction is DHNA-CoA (53, 54). The substrate of DHNA-CoA synthase, OSB-CoA, is highly unstable at physiological pH, and will spontaneously form the spirodilactone *in vitro* (Fig. 1-5; 38, 55). It is not certain that this occurs *in vivo*, nor is it known if or how OSB spirodilactone would be recycled (52). The identity of DHNA-CoA synthase in plants is still unknown, although homology searches using *E. coli* menB retrieve a very likely ortholog in *Arabidopsis*, At1g60550. Most DHNA-CoA synthases, including those from plants, cyanobacteria, and γ-proteobacteria, fall into the type I class that are dependent on the binding of bicarbonate, which is used as a catalytic base (54). The type II enzymes, found in actinobacteria and archaea, use a side-chain carboxylate from one its acidic amino acids as the catalytic base instead of bicarbonate (54).

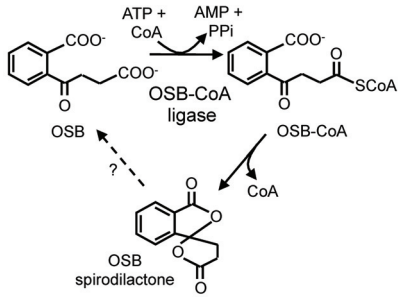


Fig. 1-5: Spontaneous hydrolysis and lactonization of OSB-CoA to the spirodilactone form at physiological pH *in vitro*. It is not known if this occurs *in vivo*, or if there are any salvaging mechanisms in place to recycle it back to OSB.

The final step in the formation of the naphthoquinone ring is the conversion of DHNA-CoA to DHNA (reaction 7, Fig. 1-4). It is one of the most misunderstood reactions in vitamin K

biosynthesis. As mentioned earlier, it has been misattributed to SHCHC synthase (MenH) and DHNA-CoA synthase (MenB) in *E. coli*. As no dedicated enzyme has been identified in any of the vitamin K-producing organisms, it has even been proposed that, like OSB-CoA, DHNA-CoA spontaneously hydrolyzes to DHNA at physiological pH (56). The focus of Chapters 2 and 3 will explore this reaction, and provide biochemical and genetic evidence that it occurs enzymatically in vitamin K-producing organisms.

Prenylation and Methylation of the Naphthoquinone Ring (reactions 8 and 9): After the naphthoquinone ring is formed, the next step in vitamin K biosynthesis is a coupling reaction with the poly-isoprenyl side chain (reaction 8, Fig. 1-4). In plants, the responsible enzyme, DHNA phytyl transferase, is an integral membrane protein encoded by *At1g60600*. This gene was identified in the *abc* (aberrant chloroplast development) T-DNA series (57). It was functionally assigned based on strong identity of its cognate enzyme with the *Synechocystis* ortholog, and the lack of phylloquinone in the *Arabidopsis* insertion mutant (57). The DHNA phytyl transferase was the first enzyme characterized in the plant phylloquinone biosynthetic pathway.

The methylation of the naphthoquinone ring is the final step in the pathway. It is catalyzed by demethylphylloquinone methyltransferase in plants (reaction 9, Fig. 1-4), which is the *At1g23360* gene product in *Arabidopsis* (28). The corresponding T-DNA insertion mutant lacks phylloquinone and accumulates the putuative substrate demethylphylloquinone (Fig. 1-4; 28). Expression of the *At1g23360* cDNA in the homologous *Synechocystis* mutant was also able to fully restore phylloquinone production.

Mutant Phenotype: In plants, the only viable phylloquinone-deficient mutant is the demethylphylloquinone methyltransferase knockout (28). This is because demethylphylloquinone is able to replace phylloquinone in photosystem I, albeit with reduced photosynthetic efficiency (28). All of the other phylloquinone biosynthetic *Arabidopsis* mutants lack demethylphylloquinone and phylloquinone, and thus display loss of photoautotrophy and seedling lethality as a result of the complete disruption of photosystem I (29, 52, 57). These mutants can be grown *in vitro* under

low light on media supplemented with sucrose, but are only able to put on a few pale-green leaves before dying. On the other hand, green algal and cyanobacterial phylloquinone-deficient mutants are able to grow photoautotrophically due to an ability to recruit plastoquinone to the photosystem I (21, 58). Since facultative anaerobic bacteria, such as *E. coli*, use menaquinone for respiration in the absence of oxygen, their growth is unaffected under aerobic conditions.

Subcellular Localization of Phylloquinone Biosynthesis in Plants: Before any of the phylloquinone biosynthetic enzymes were identified in plants, radiolabeling studies established decades ago that the prenylation and methylation reactions occur in the chloroplast envelope and thylakoids, respectively (41, 42, 59). These observations are consistent with the plastidial functions of phylloquinone discussed earlier. It was just within the last few years confirmed that both *Arabidopsis* DHNA phytyl transferase and demethylphylloquinone methyltransferase contain N-terminal signaling peptides that target the enzymes to plastids (28, 57). The same has been found for isochorismate synthases 1 and 2 (46, 47), PHYLLO (29), and OSB-CoA ligase (52). Intriguingly, however, in most monocot and dicot species, OSB-CoA ligase also displays a C-terminal tripeptide characteristic of a peroxisomal targeting signal type 1 (60). By fusing an N-terminal fluorescent protein to the *Arabidopsis* OSB-CoA ligase (AAE14), it was shown in onion epidermal cells that the protein is targeted to peroxisomes (60). It is important to point out, however, that by fusing fluorescent proteins to the N- or C-terminal ends of AAE14 that a bias is introduced into these experiments that masks the plastidial and peroxisomal targeting signals, respectively. Thus, in order to truly sort out if AAE14 is localized to the plastid, peroxisome, or even both organelles, the experiment that needs to be done is inserting a fluorescent protein between the N- and C-terminal ends of the protein.

Compounding the mystery surrounding the subcellular targeting of OSB-CoA ligase is the localization of the predicted DHNA-CoA synthase in plants (recall that this enzyme uses the product of OSB-CoA ligase as substrate). Proteomic studies have found signature peptides of the putative *Arabidopsis* and spinach DHNA-CoA synthases in highly-purified peroxisomes (60, 61). These enzymes contain N-terminal extensions bearing predicted peroxisomal targeting

signals type 2, and by fusing a C-terminal fluorescent protein to the *Arabidopsis* enzyme, its peroxisomal localization was confirmed in onion epidermal cells (60, 61). Interestingly, DHNA-CoA synthase orthologs in certain green and red algae, mosses, and *Paulinella chromatophora* lack peroxisomal targeting signals. Thus, suggesting that a possible split between plastids and peroxisomes is unique to the higher plant lineages. The discovery of two *Arabidopsis* DHNA-CoA thioesterases described in Chapter 3 will shed more light on the organellar split of phylloquinone biosynthesis in plants.

Evolution of Vitamin K Biosynthesis in Photosynthetic Eukaryotes: As described earlier, menaquinones are the ancestral quinone, having evolved more than 2.5 billion years ago (18). Most of the existing prokaryotic organisms still retain the ability to synthesize vitamin K (17). Non-photosynthetic eukaryotes are unable to produce vitamin K, so those that require it must obtain it exogenously. In contrast, those organisms in the eukaryotic lineages stemming from the plastid-engulfment events gained the capacity to synthesize vitamin K. In the more recent event, the cercozoan amoeboid *P. chromatophora* acquired two kidney-shaped cyanobacterial endosybionts similar to *Synechococcus* called the chromatophores (reviewed in 62). It appears that *P. chromatophora* is able to make vitamin K using biosynthetic genes encoded within the chromatophore genome. In an older separate event, often referred to as primary plastid endosymbiosis, plants and algae captured genes to synthesize vitamin K, which have been retained in subsequent engulfments (*e.g.* diatoms; Fig. 1-6). Plants and green algae synthesize phylloquinone using nuclear-encoded genes (Fig. 1-6). In contrast, the orthologs in *Cyanidales* (menaquinone-synthesizing red algae), except for *menG*, are clustered within the plastid genome—likely reminiscent of the cyanobacterial endosymbiont (Fig. 1-6). In diatoms, which arose from the engulfment a red algal in a secondary endosymbiotic event, the vitamin K biosynthetic genes are encoded in the nucleus (Fig. 1-6).

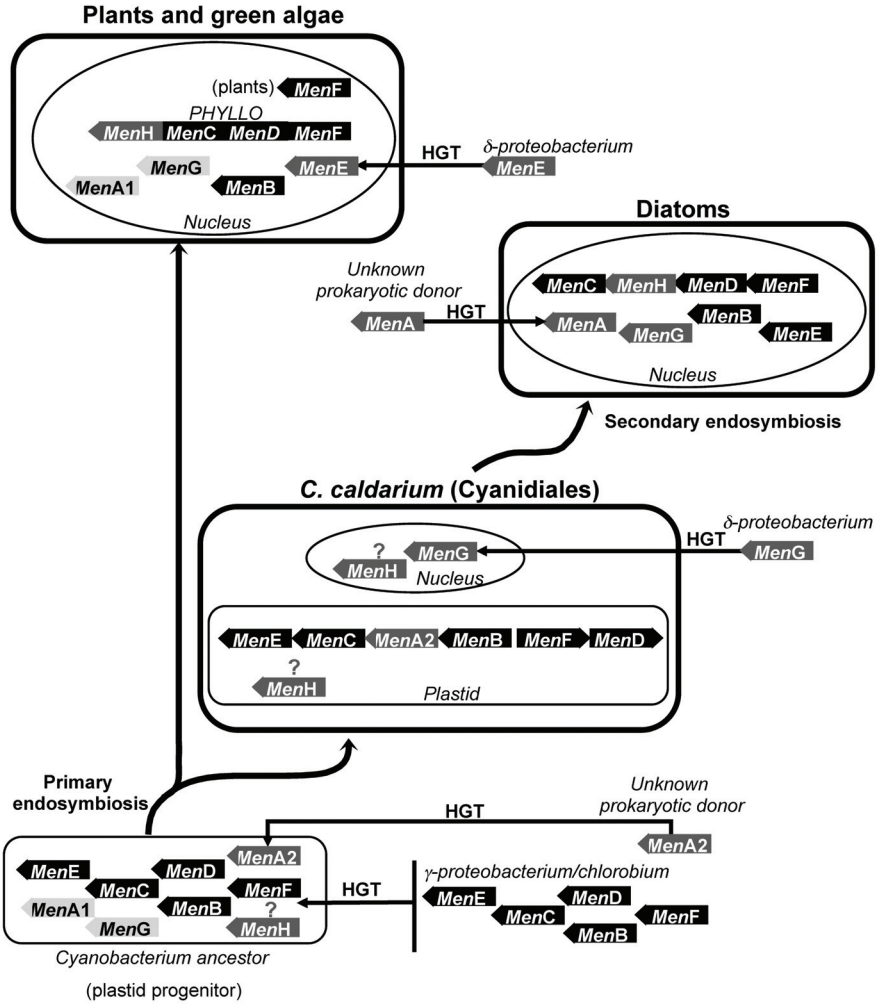


Fig. 1-6: Proposed evolution of the *men* genes in photosynthetic eukaryotes by Gross *et al.* (63). Arrows symbolize gene transfers. Note that *menH* homologs are not detected in the plastid genomes of *Cyanidiales*. As for *men*A, the cyanobacterial progenitor of plastids would have harbored two cognate homologs: *men*A1, of cyanobacterial descent, and *men*A2, acquired by HGT from an unknown prokaryotic donor. *Cyanidiales* would have lost *men*A1 and retained *men*A2; the opposite would have happened in plants and green algae. HGT, Horizontal Gene Transfer.

Phylogenetic reconstructions have revealed that the *menF,D,C,E,B* orthologs in photosynthetic eukaryotes are more closely related to those from chlorobi and γ-proteobacteria than cyanobacteria (63). Such a finding appears to contradict the theory that the ability of photosynthetic eukaryotes to synthesize vitamin K came with the engulfment of the plastid. To reconcile this observation, it has been proposed that the *men* genes in the cyanobacterial progenitor of plastids were obtained through events of horizontal gene transfer (HGT) prior to endosymbiosis (Fig. 1-6; 63). This chimeric relic now "survives" as a *men* gene cluster in the

plastid genome of extant *Cyanidiales,* and in the nuclear genomes of diatoms, green algae and plants (63).

It appears that in green algae and plants the *menG* ortholog is of cyanobacterial origin, and was likely present in the plastid ancestor (63). In *Cyanidales* and diatoms, however, the *menG* orthologs are not of cyanobacterial origin, but instead appear to have been obtained after plastid engulfment through HGT with a δ-proteobacterium (Fig. 1-6; 63). For *menA*, it is proposed that the plastid ancestor had two types, one of cyanobacterial ancestry and one coming from an unidentified prokaryotic donor (63). Green algae and plants retained the cyanobacterial type, while red algae kept the second type (Fig. 1-6; 63). It is then suggested that diatoms obtained a third type from an unknown source (Fig. 1-6; 63). Complicating the picture even more, it is believed that green algae and plants obtained their *menE* genes directly from a δ-proteobacterium (Fig. 1-6; 63). All in all, these reconstructions have illustrated that the origin of vitamin K biosynthesis in photosynthetic eukaryotes is complex and chimeric. In Chapter 5 I will discuss using comparative genomics how the discovery of DHNA-CoA thioesterases questions this evolutionary model by Gross *et al.* (63).

Engineering of Phylloquinone Biosynthesis in Plants

The only data for engineering of phylloquinone biosynthesis in plants comes from salicylate studies in tobacco and functional complementation experiments in *Arabidopsis*. Over-expression of the *E. coli* isochorismate synthase in tobacco plastids led to a four-fold increase in phylloquinone (48). On the other hand, overexpression of OSB-CoA ligase and demethylphylloquinone methyltransferase had no impact on phylloquinone content (28, 52). In Chapter 4 I will report the results from a metabolic engineering study designed to overexpress the peroxisomal steps of phylloquinone biosynthesis in plastids of *Camelina sativa*.

References

1. Dam H (1929) Cholesterinositoffweschel in huhnereiern und huhnchen . *Biochememische Zeitschrift* 215, 475-492.
2. McFarlane, W.D., Graham, W.R. Jr., Hall, G.E. (1931) The influence of different protein concentrates on the growth of baby chicks, when fed as the source of protein in various simplified diet. *The Journal of Nutrition* 4, 331-349.

3. Holst WF & Halbrook ER (1933) A "scurvy-like" disease in chicks. *Science* 77, 354.
4. Dam H & Schonheyder F (1934) A deficiency disease in chicks resembling scurvy. *Biochem J* 28, 1355-1359.
5. Almquist HJ & Stokstad ELR (1935) Hemorrhagic chick disease of dietary origin. *Journal of Biological Chemistry}* 111, 105-113.
6. Schønheyder F (1935) The anti-hemorrhagic vitamin of the chick. measurement and biological action. *Nature* 136, 653.
7. Dam H (1935) The antihaemorrhagic vitamin of the chick. *Biochem J* 29, 1273-1285.
8. Binkley SB, MacCorquodale DW, Thayer SA & Doisy EA (1939) The isolation of vitamin K_1. *Journal of Biological Chemistry}* 130, 219-234.
9. MacCorquodale DW, Binkley SB, Thayer SA & Doisy EA (1939) ON THE CONSTITUTION OF VITAMIN K1. *J Am Chem Soc* 61, 1928-1929.
10. MacCorquodale DW*, et al* (1939) The constitution and synthesis of vitamin K_1. *Journal of Biological Chemistry}* 131, 357-370.
11. McKee RW, Binkley SB, MacCorquodale DW, Thayer SA & Doisy EA (1939) THE ISOLATION OF VITAMINS K1 AND K2. *J Am Chem Soc* 61, 1295-1295.
12. Doisy EA, Binkley SB & Thayer SA (1941) Vitamin K. *Chem Rev* 28, 477-517.
13. Oostende C, Widhalm JR & Basset GJ (2008) Detection and quantification of vitamin K(1) quinol in leaf tissues. *Phytochemistry* 69, 2457-2462.
14. Widhalm JR, van Oostende C, Furt F & Basset GJ (2009) A dedicated thioesterase of the hotdog-fold family is required for the biosynthesis of the naphthoquinone ring of vitamin K1. *Proc Natl Acad Sci U S A* 106, 5599-5603.
15. Furt F*, et al* (2010) A bimodular oxidoreductase mediates the specific reduction of phylloquinone (vitamin K) in chloroplasts. *Plant J* 64, 38-46.
16. Chu PH, Huang TY, Williams J & Stafford DW (2006) Purified vitamin K epoxide reductase alone is sufficient for conversion of vitamin K epoxide to vitamin K and vitamin K to vitamin KH2. *Proc Natl Acad Sci U S A* 103, 19308-19313.
17. Collins MD & Jones D (1981) Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implication. *Microbiol Rev* 45, 316-354.
18. Schoepp-Cothenet B*, et al* (2009) Menaquinone as pool quinone in a purple bacterium. *Proc Natl Acad Sci U S A* 106, 8549-8554.
19. Nowicka B & Kruk J (2010) Occurrence, biosynthesis and function of isoprenoid quinones. *Biochim Biophys Acta* 1797, 1587-1605.
20. Richardson DJ (2000) Bacterial respiration: A flexible process for a changing environment. *Microbiology* 146 (Pt 3), 551-571.
21. Lefebvre-Legendre L*, et al* (2007) Loss of phylloquinone in *Chlamydomonas* affects plastoquinone pool size and photosystem II synthesis. *J Biol Chem* 282, 13250-13263.
22. Brettel K, Sétif P & Mathis P (1986) Flash-induced absorption changes in photosystem I at low temperature: Evidence that the electron acceptor A1 is vitamin K1. *FEBS Lett* 203, 220.
23. Petersen J, Stehlik D, Gast P & Thurnauer M (1987) Comparison of the electron spin polarized spectrum found in plant photosystem I and in iron-depleted bacterial reaction centers with time-resolved K-band EPR; evidence that the photosystem I acceptor A_1 is a quinone. *Photosynthesis Res* 14, 15-30.
24. Sigfridsson K, Hansson O & Brzezinski P (1995) Electrogenic light reactions in photosystem I: Resolution of electron-transfer rates between the iron-sulfur centers. *Proc Natl Acad Sci U S A* 92, 3458-3462.
25. Mimuro M*, et al* (2005) The secondary electron acceptor of photosystem I in *Gloeobacter violaceus* PCC 7421 is menaquinone-4 that is synthesized by a unique but unknown pathway. *FEBS Lett* 579, 3493-3496.
26. Yoshida E, Nakamura A & Watanabe T (2003) Reversed-phase HPLC determination of chlorophyll a' and naphthoquinones in photosystem I of red algae: Existence of two menaquinone-4 molecules in photosystem I of *Cyanidium caldarium. Anal Sci* 19, 1001-1005.

27. Ikeda Y*, et al* (2008) Photosystem I complexes associated with fucoxanthin-chlorophyll-binding proteins from a marine centric diatom, *Chaetoceros gracilis. Biochim Biophys Acta* 1777, 351-361.
28. Lohmann A*, et al* (2006) Deficiency in phylloquinone (vitamin K1) methylation affects prenyl quinone distribution, photosystem I abundance, and anthocyanin accumulation in the *Arabidopsis* AtmenG mutant*. J Biol Chem* 281, 40461-40472.
29. Gross J*, et al* (2006) A plant locus essential for phylloquinone (vitamin K1) biosynthesis originated from a fusion of four eubacterial genes*. J Biol Chem* 281, 17189-17196.
30. Goodstadt L & Ponting CP (2004) Vitamin K epoxide reductase: Homology, active site and catalytic mechanism*. Trends Biochem Sci* 29, 289-292.
31. Singh AK, Bhattacharyya-Pakrasi M & Pakrasi HB (2008) Identification of an atypical membrane protein involved in the formation of protein disulfide bonds in oxygenic photosynthetic organisms*. J Biol Chem* 283, 15762-15770.
32. Li T*, et al* (2004) Identification of the gene for vitamin K epoxide reductase*. Nature* 427, 541-544.
33. Li W*, et al* (2010) Structure of a bacterial homologue of vitamin K epoxide reductase*. Nature* 463, 507-512.
34. Booth SL & Suttie JW (1998) Dietary intake and adequacy of vitamin K*. J Nutr* 128, 785-788.
35. Hiratsuka T*, et al* (2008) An alternative menaquinone biosynthetic pathway operating in microorganisms*. Science* 321, 1670-1673.
36. Young IG (1975) Biosynthesis of bacterial menaquinones. menaquinone mutants of *Escherichia coli. Biochemistry* 14, 399-406.
37. Dansette P & Azerad R (1970) A new intermediate in naphthoquinone and menaquinone biosynthesis*. Biochem Biophys Res Commun* 40, 1090-1095.
38. Heide L, Kolkmann R, Arendt S & Leistner E (1982) Enzymic synthesis of o-succinylbenzoyl-CoA in cell-free extracts of anthraquinone producing *Galium mollugo* L. cell suspension cultures*. Plant Cell Rep* 1, 180-182.
39. Thomas G & Threlfall DR (1974) Incorporation of shikimate and 4-(2'-carboxyphenyl)-4-oxobutyrate into phylloquinone*. Phytochemistry* 13, 807.
40. Hutson KG & Threlfall DR (1980) Asymmetric incorporation of 4-(2'-carboxyphenyl)-4-oxobutyrate into phylloquinone by *Zea mays. Phytochemistry* 19, 535.
41. Schultz G, SOLL J & Ellerbrock BH (1981) Site of prenylation reaction in synthesis of phylloquinone (vitamin K1) by spinach chloroplasts*. Eur J Biochem* 117, 329-332.
42. Gaudillière J, d'Harlingue A, Camara B & Monéger R (1984) Prenylation and methylation reactions in phylloquinone (vitamin K_1) synthesis in *Capsicum annuum* plastids*. Plant Cell Rep* 3, 240-242.
43. Meganathan R (2001) Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): A perspective on enzymatic mechanisms*. Vitam Horm* 61, 173-218.
44. Lee PT, Hsu AY, Ha HT & Clarke CF (1997) A C-methyltransferase involved in both ubiquinone and menaquinone biosynthesis: Isolation and identification of the *Escherichia coli* ubiE gene*. J Bacteriol* 179, 1748-1754.
45. Wildermuth MC, Dewdney J, Wu G & Ausubel FM (2001) Isochorismate synthase is required to synthesize salicylic acid for plant defense*. Nature* 414, 562-565.
46. Strawn MA*, et al* (2007) *Arabidopsis* isochorismate synthase functional in pathogen-induced salicylate biosynthesis exhibits properties consistent with a role in diverse stress responses*. J Biol Chem* 282, 5919-5933.
47. Garcion C*, et al* (2008) Characterization and biological function of the ISOCHORISMATE SYNTHASE2 gene of *Arabidopsis. Plant Physiol* 147, 1279-1287.
48. Verberne MC, Sansuk K, Bol JF, Linthorst HJ & Verpoorte R (2007) Vitamin K1 accumulation in tobacco plants overexpressing bacterial genes involved in the biosynthesis of salicylic acid*. J Biotechnol* 128, 72-79.
49. Jiang M*, et al* (2007) Menaquinone biosynthesis in *Escherichia coli:* Identification of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate as a novel intermediate and re-evaluation of MenD activity*. Biochemistry* 46, 10979-10989.

50. Jiang M*, et al* (2008) Identification and characterization of (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of *Escherichia coli. Biochemistry* 47, 3426-3434.
51. Kolkmann R & Leistner E (1987) 4-(2'-carboxyphenyl)-4-oxobutyryl coenzyme A ester, an intermediate in vitamin K2 (menaquinone) biosynthesis. *Z Naturforsch C* 42, 1207-1214.
52. Kim HU, van Oostende C, Basset GJ & Browse J (2008) The AAE14 gene encodes the *Arabidopsis* o-succinylbenzoyl-CoA ligase that is essential for phylloquinone synthesis and photosystem-I function. *Plant J* 54, 272-283.
53. Truglio JJ*, et al* (2003) Crystal structure of *Mycobacterium tuberculosis* MenB, a key enzyme in vitamin K2 biosynthesis. *J Biol Chem* 278, 42352-42360.
54. Jiang M, Chen M, Guo Z & Guo Z (2010) A bicarbonate cofactor modulates 1,4-dihydroxy-2-naphthoyl-coenzyme A synthase in menaquinone biosynthesis of *Escherichia coli. Journal of Biological Chemistry}* 285, 30159-30169.
55. Meganathan R & Bentley R (1979) Menaquinone (vitamin K2) biosynthesis: Conversion of o-succinylbenzoic acid to 1,4-dihydroxy-2-naphthoic acid by *Mycobacterium phlei* enzymes. *J Bacteriol* 140, 92-98.
56. Sakuragi Y & Bryant DA (2006) Genetic manipulation of quinone biosynthesis in cyanobacteria. In JH Golbeck, ed, Photosystem I: The Plastocyanin:Ferredoxin Oxidoreductase. Advances in Photosynthesis and Respiration. Vol 24 in ed Golbeck JH (Springer, Dordrecht, The Netherlands), pp 205-222.
57. Shimada H*, et al* (2005) Inactivation and deficiency of core proteins of photosystems I and II caused by genetical phylloquinone and plastoquinone deficiency but retained lamellar structure in a T-DNA mutant of *Arabidopsis. Plant J* 41, 627-637.
58. Johnson TW*, et al* (2000) Recruitment of a foreign quinone into the A(1) site of photosystem I. I. genetic and physiological characterization of phylloquinone biosynthetic pathway mutants in *Synechocystis* sp. pcc 6803. *J Biol Chem* 275, 8523-8530.
59. Kaiping S, Soll J & Schultz G (1984) Site of methylation of 2-phytyl-1,4-naphthoquinol in phylloquinone (vitamin K1) synthesis in spinach chloroplasts. *Phytochemistry* 23, 89.
60. Babujee L*, et al* (2010) The proteome map of spinach leaf peroxisomes indicates partial compartmentalization of phylloquinone (vitamin K1) biosynthesis in plant peroxisomes. *J Exp Bot* 61, 1441-1453.
61. Reumann S*, et al* (2007) Proteome analysis of *Arabidopsis* leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19, 3170-3193.
62. Keeling PJ (2004) Diversity and evolutionary history of plastids and their hosts. *Am J Bot* 91, 1481-1493.
63. Gross J, Meurer J & Bhattacharya D (2008) Evidence of a chimeric genome in the cyanobacterial ancestor of plastids. *BMC Evol Biol* 8, 117.

CHAPTER 2

Discovery of a Missing Enzyme in Vitamin K Biosynthesis: Characterization of a Cyanobacterial 1,4-dihydroxy-2-naphthoyl-CoA Thioesterase

Preface

When I first began working on my Ph.D. one of the predominant unanswered questions in vitamink K biosynthesis centered around how 1,4-dihydroxy-2-napthoate (DHNA) was formed from its CoA-thioester (DHNA-CoA; reaction 7, Fig. 1-4). This is the final step in the production of the naphthoquinone ring moiety of vitamin K and is cardinal to permit the conjugation of the polyisoprenyl side chain. Previous studies had proposed that two other enzymes in the vitamin K biosynthetic pathway were responsible for the hydrolysis of DHNA-CoA to form DHNA, but were later shown to catalyze upstream reactions instead. It was even suggested that this reaction occurred spontaneously, thus requiring no enzyme. It was at this juncture that we began exploring the possibility that an unknown enzyme remained to be discovered in vitamin K biosynthesis. Using a multi-faceted approach consisting of comparative genomics, protein biochemistry, genetics, and metabolic profiling, we were able to identify and characterize a genuine "missing" enzyme in the vitamin K biosynthetic pathway. This achievement is presented here in Chapter 2 essentially as it appears in our manuscript published in *Proceedings of the National Academy of Sciences of the United States of America*:

A dedicated thioesterase of the Hotdog-fold family is required for the biosynthesis of the naphthoquinone ring of vitamin K1

Joshua R. Widhalm, Chloë van Oostende, Fabienne Furt and Gilles J. C. Basset

Published on April 7, 2009 in volume 106 (no. 14) on pages 5599-5603 of *Proceedings of the National Academy of Sciences of the United States of America*

Abstract

Phylloquinone (vitamin K_1) is a bipartite molecule that consists of a naphthoquinone ring attached to a phytyl side chain. The coupling of these 2 moieties depends on the hydrolysis of the CoA thioester of 1,4-dihydroxy-2-naphthoate (DHNA), which forms the naphthalenoid backbone. It is not known whether such a hydrolysis is enzymatic or chemical. In this study, comparative genomic analyses identified orthologous genes of unknown function that in most species of cyanobacteria cluster with predicted phylloquinone biosynthetic genes. The encoded approximately 16-kDa proteins display homology with some Hotdog domain-containing CoA thioesterases that are involved in the catabolism of 4-hydroxybenzoyl-CoA and gentisyl-CoA (2,5-dihydroxybenzoyl-CoA) in certain soil-dwelling bacteria. The *Synechocystis* ortholog, encoded by gene slr0204, was expressed as a recombinant protein and was found to form DHNA as reaction product. Unlike its homologs in the Hotdog domain family, Slr0204 showed strict substrate specificity. The *Synechocystis slr0204* knockout was devoid of DHNA-CoA thioesterease activity and accumulated DHNA-CoA. As a result, knockout cells contained less than 8% of the phylloquinone in their wild-type counterparts and displayed the typical photosensitivity to high light associated to phylloquinone deficiency in cyanobacteria.

Introduction

Vitamin K encompasses a class of fat-soluble compounds formed from a naphthoquinone ring attached to a polyisoprenyl chain of variable length and saturation. Its main natural forms are vitamin K_1 (phylloquinone), which contains a partially saturated C-20 phytyl side chain, and vitamin K_2 (menaquinone), whose usually longer side chain is fully conjugated (Fig. 2-1). Phylloquinone is synthesized by plants and most cyanobacteria (1), whereas menaquinone is synthesized by facultative anaerobic bacteria, archaea, and some cyanobacteria (1, 2). Vitamin K plays very different roles in the organisms that synthesize it compared to those that do not. Facultative anaerobic bacteria use menaquinone as an electron transporter in their respiratory chain (3). Similarly, in cyanobacteria and plants, phylloquinone serves as the one-electron carrier at the A1 site of photosystem I (4). In contrast, vertebrates, which do not synthesize vitamin K,

use it as a cofactor for certain carboxylases involved in blood coagulation, bone and vascular metabolism, and cell cycle regulation (5). In mammals, vitamin K is also known to act as a signaling molecule that regulates the transcription of genes involved in its own metabolism (6).

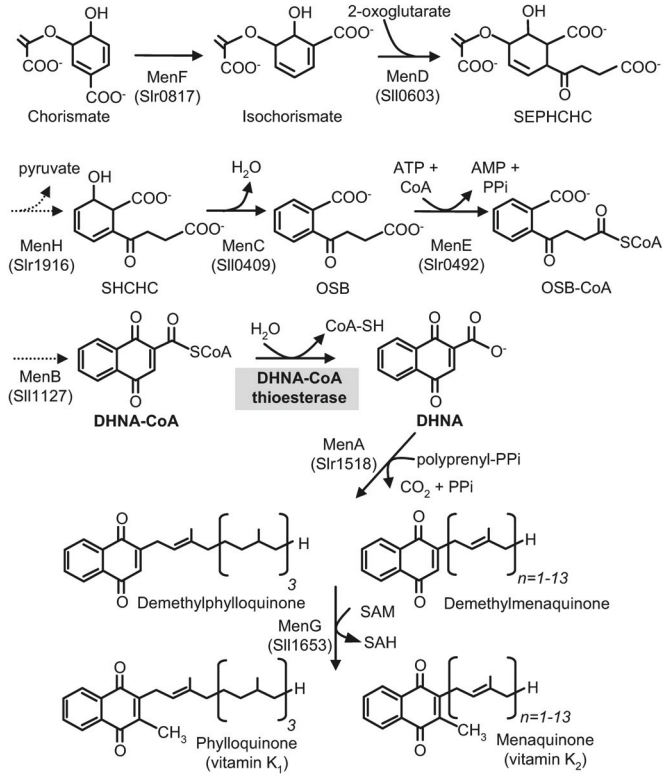


Fig. 2-1. The biosynthetic pathway of vitamin K in *E. coli* and *Synechocystis* sp. PCC 6803. The *E. coli* Men enzymes and the protein number of their orthologs in *Synechocystis* (brackets) corresponding to each known step are indicated. MenF, isochorismate synthase; MenD, SEPHCHC synthase; MenH, SHCHC synthase (note that this enzyme is currently mis-annotated in most databases as DHNA-CoA thioesterase); MenC, OSB synthase; MenE, OSB-CoA ligase; MenB, DHNA synthase; MenA, DHNA prenyltransferase; MenG, demethyl-menaquinone/demethyl-phylloquinone methyltransferase. SEPHCHC, 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylic acid; SHCHC, (1*R*,6*R*)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylic acid.

Depending on the organisms, there are two different routes for the biosynthesis of the naphthoquinone moiety of vitamin K. Both start from chorismate. In the first pathway, that is specific to certain menaquinone-synthesizing microorganisms, the naphthalenoid backbone originates from a chorismate-inosine conjugate called futalosine (7). In the second pathway, which includes all of the phylloquinone producers, the core of the naphthoquinone ring comes from 1,4-dihydroxy-2-naphthoate (DHNA). It is subsequently coupled to the polyisoprenyl side

chain and then methylated (Fig. 2-1). The DHNA-pathway was initially described in *Escherichia coli* (8), and later using genetic approaches in cyanobacteria (9–11) and plants (12–16). In this pathway, chorismate is first converted into isochorismate, to which a succinyl side chain is added at the C2 position. After elimination of pyruvate and aromatization of the cyclohexadiene ring, the succinyl chain is activated by ligation with CoA, and then cyclized, yielding DHNA-CoA (Fig. 2-1). The subsequent removal of CoA is of special interest, for it frees the carboxyl group of DHNA and allows its conjugation to the polyisoprenyl chain for anchoring into biological membranes. The hydrolysis of DHNA-CoA has long been a puzzling mystery of vitamin K biosynthesis, and it is actually not clear whether this step is enzyme-driven or a purely chemical process (1). It was first thought to occur concomitantly to the cyclization step catalyzed by DHNA synthase, hence the name of this enzyme (17). Such a mechanism was later invalidated when it was shown that the product of the DHNA synthase-catalyzed reaction was not DHNA but DHNA-CoA (18). The identification in the menaquinone biosynthetic operon of *E. coli* of a gene encoding for a protein termed MenH (8), whose α/β-fold type 1 motif is commonly found among thioesterases compounded the confusion. Despite the lack of direct evidence that MenH or its orthologs displayed some activity against DHNA-CoA, the protein was assumed to correspond to the missing CoA thioesterase (8). It is not until recently that MenH was shown to lack thioesterase activity and to catalyze in fact an earlier step in the pathway (Fig. 2-1) (19), reopening *de facto* the search for a putative DHNA-CoA thioesterase.

In this study, we used comparative genomics to infer the existence of DHNA-CoA thioesterase in the phylloquinone biosynthetic pathway of cyanobacteria. We provide biochemical and genetic evidence that the encoding gene is required for phylloquinone biosynthesis.

Results

Cyanobacteria Contain Putative CoA Thioesterases That Cluster with Predicted Phylloquinone Biosynthetic Genes: Because prokaryotic genes that are involved in the same cellular process tend to cluster together (20), we used the online database for comparative genomics SEED and its tools (http://theseed.uchicago.edu/FIG/index.cgi) to search for conserved

physical associations between known vitamin K biosynthetic genes and genes whose function is unknown or *a priori* unrelated. These *in silico* searches detected a group of cyanobacterial orthologs that cluster with the near complete set of phylloquinone biosynthetic genes (Fig. 2-2). The encoded proteins (approximately 16-kDa) have unknown function, but feature a conserved domain (cd03440), named Hotdog, that is found in certain thioesterases (21). Consistent with this, the cyanobacterial proteins share 17–25% identity with *Pseudomonas sp. strain CBS-3* 4-hydroxybenzoyl-CoA thioesterase (4-HBT; EC 3.1.2.23) and 16%–23% identity with *Bacillus halodurans C-125* gentisyl-CoA (2,5-dihydroxybenzoyl-CoA) thioesterase, two members of the superfamily of Hotdog domain-containing proteins (22, 23). Because these two enzymes belong to catabolic pathways that are specific to a few clades of soil-dwelling bacteria, we hypothesized that their cyanobacterial homologs were likely to have a different function and were good candidates for the missing DHNA-CoA thioesterase of phylloquinone biosynthesis.

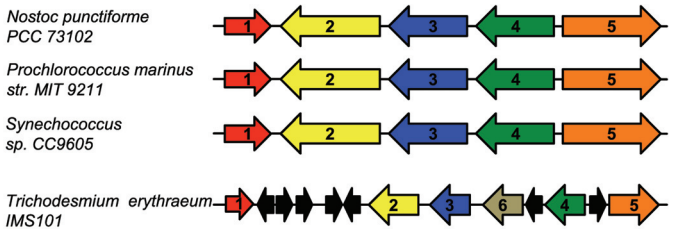


Fig. 2-2. Clustering of a putative CoA thioesterase with predicted phylloquinone biosynthetic genes in cyanobacteria. Matching colors and numbers indicate orthology. 1, putative CoA-thioesterase; 2, OSB-CoA ligase; 3, OSB synthase; 4, DHNA prenyltransferase; 5, isochorismate synthase; 6, SHCHC synthase. Black arrows indicate genes of unknown function or whose function is a priori not directly connected to phylloquinone biosynthesis. Note that the gene cluster in *T. erythraeum* is not to scale with that of the other species.

***Synechocystis* Protein Slr0204 Displays Highly Specific DHNA-CoA Thioesterase Activity:** BLASTp searches using the protein sequences of the candidates previously identified in SEED detected a single ortholog encoded by gene slr0204 in *Synechocystis* sp. PCC 6803. To determine whether the corresponding protein bore DHNA-CoA thioesterase activity, gene slr0204 was cloned into bacterial expression vector pET43.1a, and the protein was overexpressed in *E. coli BL21 (DE3).* DHNA-CoA thioesterase activity was assayed in desalted extracts of the overexpressor and the vector-alone control by monitoring the consumption of DHNA-CoA and the

simultaneous formation of DHNA (Fig. 2-3). The hydrolysis of DHNA-CoA was readily detected in the extract of the overexpressor, but not in the vector-alone control (Fig. 2-3), providing initial biochemical evidence for DHNA-CoA thioesterase activity associated with Slr0204. No activity was detected when the extracts were boiled before the addition of the substrate (Fig. 2-3), confirming that the formation of DHNA associated with the overexpressor extract was not simply due to the chemical hydrolysis of DHNA-CoA. To corroborate this result and to test for the substrate specificity of Slr0204, the 6xhis-tagged protein was isolated by affinity chromatography. The purified enzyme hydrolyzed DHNA-CoA, but unlike its 4-HBT and gentisyl-CoA thioesterase homologs (23, 24), it did not display any activity against benzoyl-CoA and phenylacetyl-CoA (Table 2-1). Nor was any activity detected against aliphatic acyl-CoA thioesters (Table 2-1). The cyanobacterial enzyme thus appears to have stringent substrate specificity.

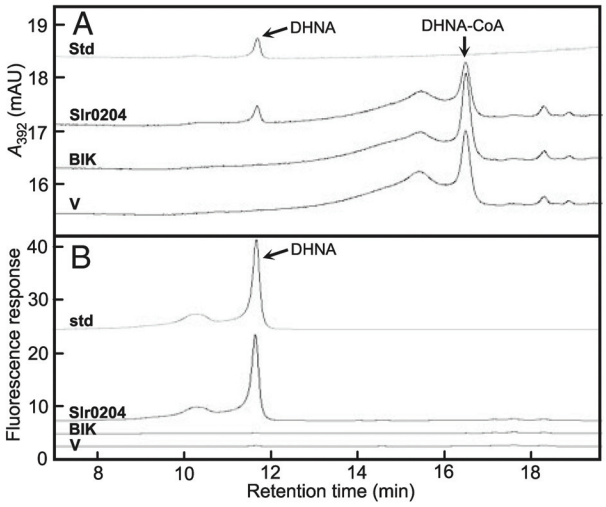


Fig. 2-3. HPLC assays of DHNA-CoA thioesterase activity in desalted extracts of *E. coli*. (*A*) Detection at A_{392} nm. (*B*) Selective detection of DHNA by fluorometry (Ex=260 nm, Em=450 nm). Extract of cells overexpressing *Synechocystis* Slr0204 protein (Slr0204), overexpressor extract boiled for 15 min before the assay (Blk), vector alone control extract (V). Assays contained 126 μM DHNA-CoA and 0.14 μg protein and were carried out for 10 min at 30 °C. The uppermost trace in each frame corresponds to 100 pmol authentic DHNA (Std).

Table 2-1. Substrate specificity of *Synechocystis* DHNA-CoA thioesterase. Purified recombinant Slr0204 (1.3–8 μg) was assayed with various CoA thioester substrates. The formation of DHNA was quantified by HPLC; for other substrates, the release of CoA was measured spectrophotometrically, using a DTNB-based dosage, as described in Materials and Methods. Data are means ± SE of three replicates.

Substrate	Specific Activity (μmol hr^{-1} mg^{-1})
DHNA-CoA (65 μM)	102 ± 14
Benzoyl-CoA (30 μM)	< 0.001
Benzoyl-CoA (120 μM)	< 0.001
Phenylacetyl-CoA (50 μM)	< 0.001
Phenylacetyl-CoA (100 μM)	< 0.001
Succinyl-CoA (14 μM)	< 0.001
Palmitoyl-CoA (11 μM)	< 0.001

***slr0204* Knockout Lacks DHNA-CoA Thioesterase Activity and Accumulates DHNA-CoA:** To investigate directly the function of slr0204 in *Synechocystis* cells, we replaced its entire sequence by a marker cassette using homologous recombination. After selection on solid media, a homogenous knockout strain was isolated and transferred to a liquid culture. Although DHNA-CoA thioesterase activity was readily detected in the desalted extracts of wild-type cells, no activity was observed in the extracts of their knockout counterparts (Fig. 2-4A). Moreover, the lack of DHNA-CoA thioesterase activity in knockout cells was paralleled by an approximately 30-fold increase of their pool size of DHNA-CoA compared to wild-type ones (Fig. 2-4B), confirming that the Slr0204 protein also functions as a DHNA-CoA thioesterase *in vivo*.

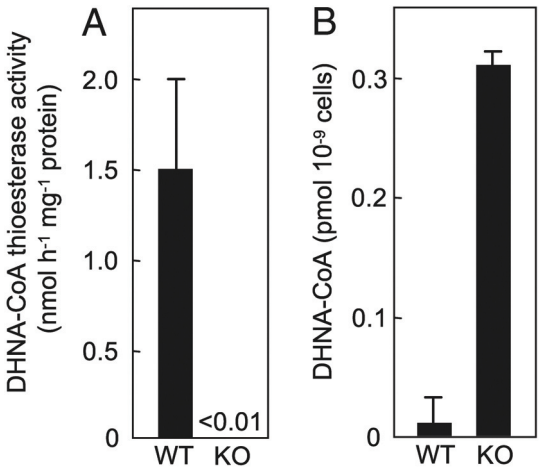


Fig. 2-4. The slr0204 mutant lacks DHNA-CoA thioesterase activity and accumulates DHNA-CoA. (*A*) Desalted extracts of wild-type (WT) and *slr0204* knockout (KO) *Synechocystis* cells were assayed for DHNA-CoA thioesterase activity for 10 min at 30 °C. Assays with wild-type and knockout extracts contained 15 and 14 μg protein, respectively. (*B*) DHNA-CoA present in the extracts of wild-type and knockout cells was quantified by HPLC-fluorescence after enzymatic conversion to DHNA with an excess of recombinant Slr0204 protein. DHNA was then extracted by phase partitioning and chemically reduced to confer fluorescent properties. The data are means of 3 replicates (DHNA-CoA thioesterase activity) and 2 replicates (DHNA-CoA level) ± SE.

***slr0204* Knockout Lacks Phylloquinone and Is Photosensitive:** As in plants, vitamin K is not the sole quinone derived from DHNA (25); we verified that Slr0204 was indeed involved in phylloquinone biosynthesis by measuring the level and redox status of this metabolite in the knockout and wild-type cells. The knockout cells were found to contain less than 8% of the phylloquinone in wild-type (Fig. 2-5A), and displayed lower phylloquinone (oxidized)/phylloquinol (reduced) ratio (Fig. 2-5A). Addition of DHNA restored the phylloquinone content of the mutant (Fig. 2-5A), whereas no significant changes were seen when *o*-succinylbenzoate (OSB)—a biosynthetic intermediate located upstream the DHNA-CoA-catalyzed reaction—was added (Fig. 2-5A). Transformation of the knockout with plasmid pSynExp2 containing the sequence of slr0204 under the control of an endogenous cyanobacterial promoter restored phylloquinone content and redox status to levels similar to that of wild-type cells (Fig. 2-5A), verifying that the lack of phylloquinone in the knockout was not caused by a secondary mutation. Compared to wild-type and complemented cells, the knockout strain displayed severe growth retardation at high light intensity, but grew normally in low light (Fig. 2-5B). This phenotype is characteristic of phylloquinone deficiency in *Synechocystis* (9).

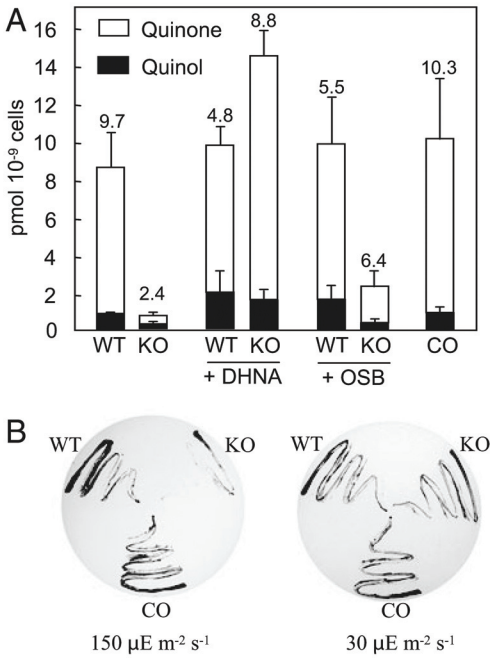


Fig. 2-5. The slr0204 mutant lacks phylloquinone and displays photosensitivity to high light intensities. (*A*) Level and redox status of phylloquinone in *Synechocystis* cells. The final concentration of DHNA and OSB in the culture media was 100 _M. Values above the bars indicate the quinone/quinol ratio. Data are means ± SE of 3 replicates. (*B*) Growth of *Synechocystis* cells in high and low light intensities. Similar numbers of cells have been plated on BG-11 medium containing glucose and no antibiotics. Plates have been incubated for 10 days (150 μE m^{-2} s^{-1}) or 14 days (30 μE m^{-2} s^{-1}) at 22°C. WT, wild type; KO, slr0204 knockout; CO, complemented mutant.

Discussion

We report here the identification of a cyanobacterial protein specifying the missing DHNA-CoA thioesterase of phylloquinone biosynthesis. This brings the total number of known phylloquinone biosynthetic genes to 9 and completes the identification of the enzymes involved in the formation of the naphthoquinone ring in cyanobacteria. Unlike their orthologs in *Nostoc punctiforme, Prochlorococcus marinus, Trichodesmium erythraeum*, and *Synechococcus*, none of the phylloquinone biosynthetic genes of *Synechocystis* sp. PCC 6803 clusters with each other. This feature actually guided our choice to work with this species, as it permitted us to knock out the DHNA-CoA thioesterase candidate without creating a negative polar effect on other phylloquinone biosynthetic genes.

Phylloquinone-synthesizing eukaryotes display homologs of cyanobacterial DHNA-CoA thioesterase (Fig. 2-6). The corresponding proteins are either plastid- or chromatophore-encoded in rhodophytes and amoeba, respectively, or nuclear-encoded in green algae and flowering plants, but then display predicted chloroplast-targeting peptides. Such a distribution is notable, for it strictly follows that of the other phylloquinone biosynthetic genes in these species, and plastids are the site of phylloquinone biosynthesis (12–15). Importantly, the rhodophytes and amoeba homologs are arranged in clusters of phylloquinone biosynthetic genes as occurs in their cyanobacterial ancestors (Fig. 2-7A). In the case of green algae and flowering plants, the genomic context of the slr0204 homologs does not permit to infer a similar functional linkage with known phylloquinone biosynthetic genes, and therefore the orthology assignment of these predicted CoA thioesterases merits further investigation.

Although slr0204 homologs are also detected in prokaryotes outside the cyanobacterial lineage, the genomic neighborhood of these genes does not indicate that they indeed encode for such an enzyme. For instance, in several facultative anaerobic bacteria, these homologs are localized in an operon-like structure comprising of components of the Tol membrane transport system and subunits of cytochrome ubiquinol oxidase (Fig. 2-7B). In our view, the identity of DHNA-CoA thioesterase remains therefore to be established in these microorganisms. There are nonetheless

two notable exceptions in proteobacteria and verrucomicrobia species, in which the slr0204 homologs cluster with predicted menaquinone biosynthetic genes (Fig. 2-7B).

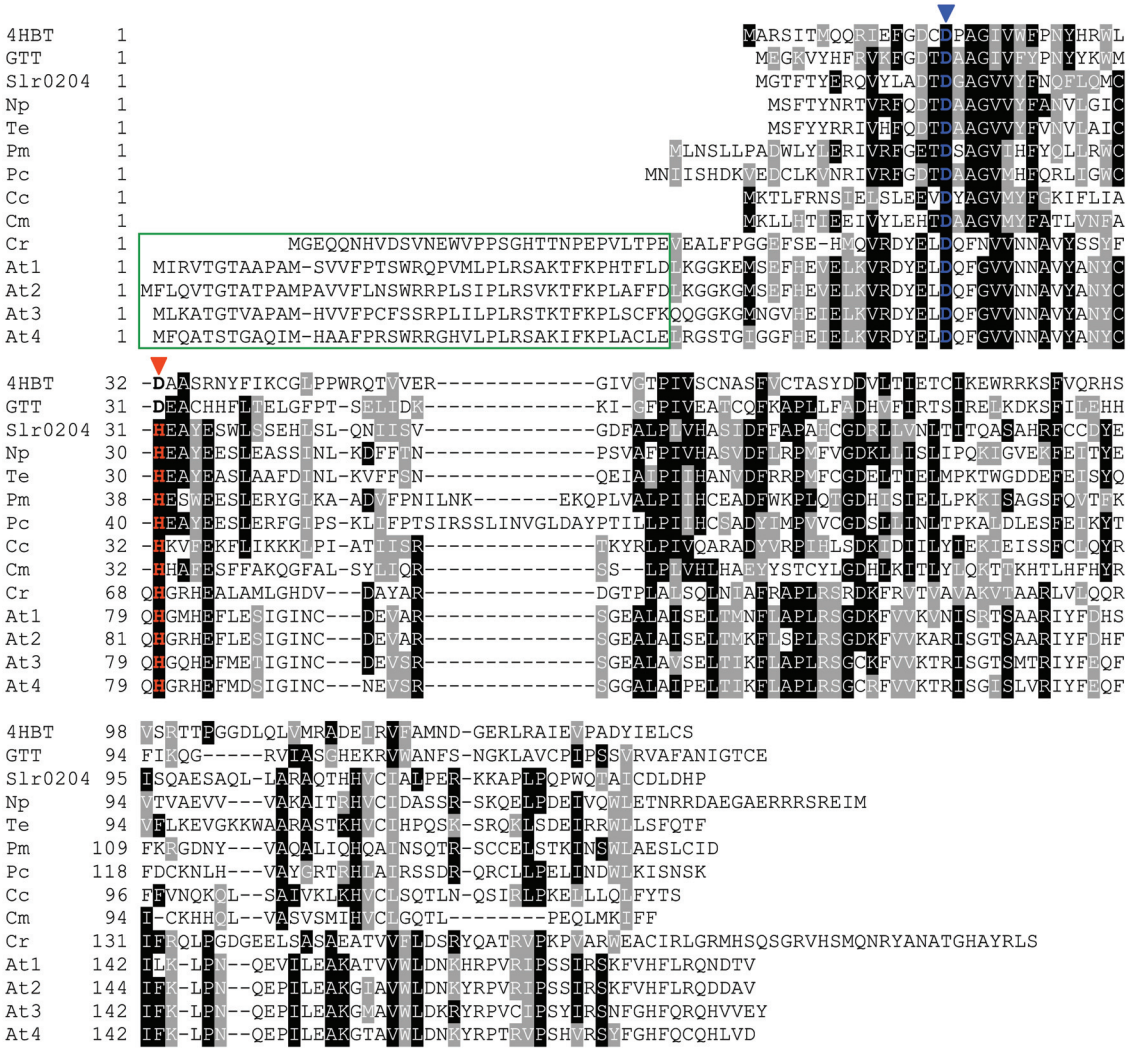


Fig. 2-6. Alignment of 4HBT proteins with Slr0204 and its homologs. Sequences are from *Pseudomonas* sp. CBS3 4-HBT (4HBT, ABQ44580), *B. halodurans* C-125 gentisyl-CoA thioesterase (GTT,NP_242865), cyanobacterial DHNA-CoA thioesterases from Chroococcales (*Synechocystis* sp. PCC 6803, Slr0204, NP_442358), Nostocales (*N. punctiforme* PCC 73102, Np, YP_001869668), Oscillatoriales (*T. erythraeum* IMS101, Te, YP_723549), and Prochlorales (*P. marinus* str. MIT9211, Pm, YP_001550076), and their eukaryotic homologs in amoeba (*P. chromatophora*, Pc, YP_002048985), rhodophytes (*C. caldarium* RK1, Cc, NP_045103; *C. merolae* 10D, Cm, NP_849030), green algae (*C. reinhardtii*, Cr, XP_001703093), and *Arabidopsis* (At1, NP_176995; At2, NP_564926; At3, NP_564457; At4, NP_174759). The *Arabidopsis* proteins are paralogs that share over 70% identity with each other. Dashes symbolize gaps introduced to maximize alignment. Identical residues are shaded in black, similar ones in gray. The blue arrowhead indicates the conserved Asp residue known to trigger the nucleophilic attack in 4-HBT. The red arrowhead points to the Asp residue that has been proposed to be important for substrate binding in 4-HBT and gentisyl-CoA thioesterase; note that the cyanobacterial and eukaryotic proteins have a His residue at this position. The N-terminal extensions that are predicted to encode plastid targeting peptides in *C. reinhardtii* and *Arabidopsis* are boxed.

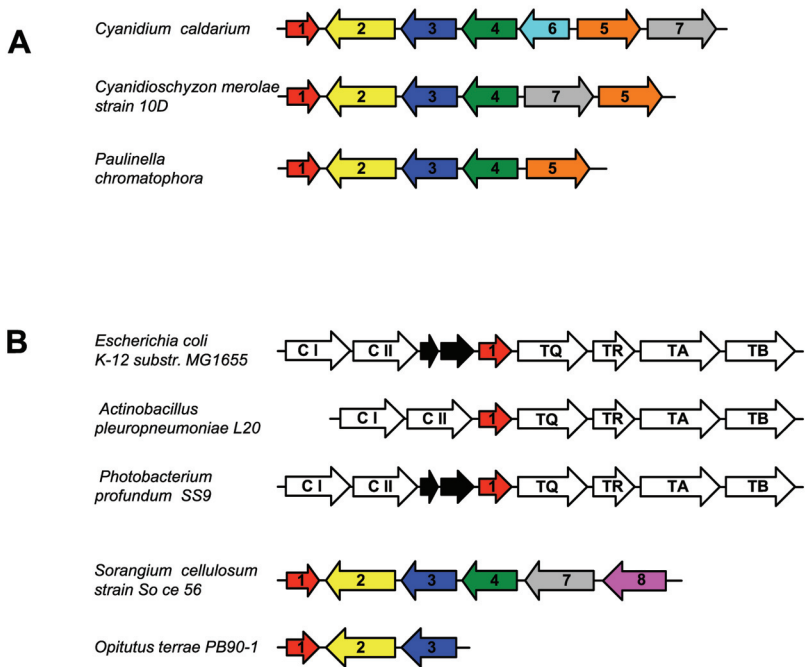
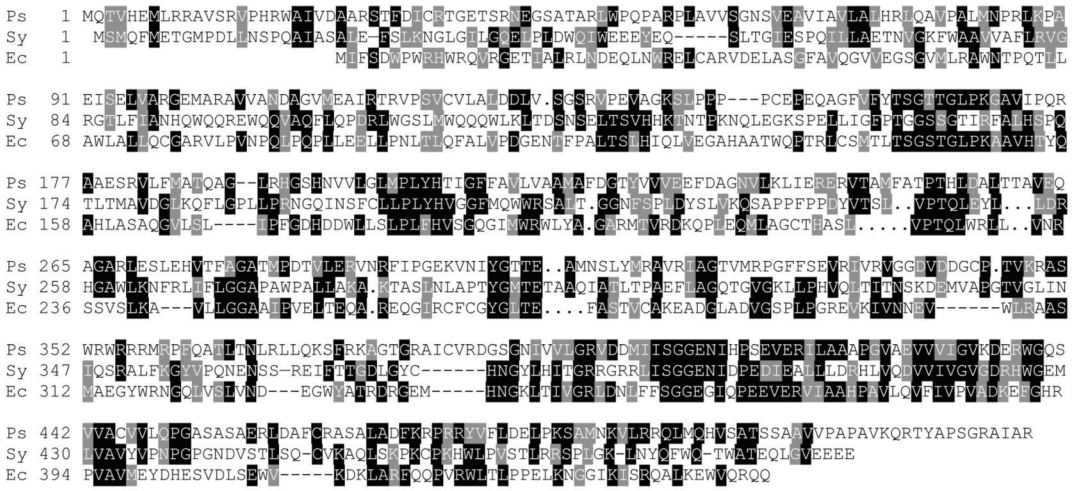


Fig. 2-7. Genomic context of the homologs of cyanobacterial DHNA-CoA thioesterase in eukaryotes and prokaryotes. (*A*) Clustering with phylloquinone biosynthetic genes in the rhodophytes *C. caldarium* and *C. merolae*, and the amoeba *P. chromatophora*. (*B*) Presence of the homolog of cyanobacterial DHNA-CoA thioesterase in the Tol operon-like structure of facultative anaerobes or in a cluster with menaquinone biosynthetic genes in the Proteobacteria *S. cellulosum* and the Verrucomicrobia *O. terrae*. 1, putative CoA-thioesterase; 2, OSB-CoA ligase; 3, OSB synthase; 4, DHNA prenyltransferase; 5, isochorismate synthase; 6, DHNA synthase; 7, SEPHCHC synthase; 8, demethylmenaquinone methyl transferase. C I and CII, cytochrome ubiquinol oxidase subunit I and II, respectively; TQ, TR, TA, and TB, Tolmembranetransport system component Q, R, A, and B, respectively. Black arrows indicate genes encoding proteins annotated as hypothetical.

The identification of DHNA-CoA thioesterase adds a functional member to the thioesterase family of Hotdog domain-containing proteins, which includes various acyl-CoA thioesterases, gentisyl-CoA thioesterase, and 4-HBT (PROSITE PF03061). DHNA-CoA thioesterase is most closely related to these last 2 members. Gentisyl-CoA thioesterase is thought to be involved in the oxidation of gentisate (2,5-dihydroxybenzoate) (23), whereas 4-HBT is part of the degradation pathway of the halogenated pollutant 4-chlorobenzoate (22). Interestingly, 4-chlorobenzoate-CoA ligase and 4-chlorobenzoyl-CoA dehalogenase, the 2 enzymes preceding 4-HBT in the degradation pathway of 4-chlorobenzoate (26), display striking homologies with OSB-CoA ligase and DHNA-synthase, respectively (Fig. 2-8). An attractive hypothesis would be that the catabolism of such a chlorinated aromatic compound has diverged from the naphthoquinone branch of vitamin K biosynthesis.

A



B

C

4-chlorobenzoate ligase	OSB-CoA ligase	
Pseudomonas	*Synechocystis*	*E. coli*
Identity/Similarity	22%/34%	20%/34%

4-chlorobenzoate dehalogenase	DHNA synthase	
Pseudomonas	*Synechocystis*	*E. coli*
Identity/Similarity	28%/44%	24%/35%

Fig. 2-8. Sequence comparisons of *Pseudomonas* sp. CBS3 4-chlorobenzoate-CoA ligase and 4-chlorobenzoate dehalogenase with OSB-CoA ligase and DHNA synthase, respectively. (*A*) *Pseudomonas* sp. CBS3 4-chlorobenzoate-CoA ligase (Ps, B42560), *Synechocystis* sp. PCC 6803 OSB-CoA ligase (Sy, NP_442252), *E. coli* OSB-CoA ligase (Ec, NP_416763). (*B*) *Pseudomonas* sp. CBS3 4-chlorobenzoate dehalogenase (Ps, A42560), *Synechocystis* sp. PCC 6803 DHNA synthase (Sy, NP_440855), *E. coli*DHNAsynthase (Ec, NP_416765). Dashes symbolize gaps introduced to maximize alignment. Identical residues are shaded in black, similar ones in gray. (*C*) Levels of sequence homology between 4chlorobenzoate-CoA ligase and OSB-CoA ligase, and between 4-chlorobenzoate dehalogenase and DHNA
synthase.

The absence of detectable activity of DHNA-CoA thioesterase against benzoyl-CoA and phenylacetyl-CoA, whereas its homologs, 4-HBT and gentisyl-CoA thioesterase, show some activity with these substrates (23, 24), is of much interest from the standpoint of the structural determinants of substrate recognition in this enzyme family. Indeed, alignment of cyanobacterial and plant DHNA-CoA thioesterases with 4-HBT and gentisyl-CoA thioesterase, reveals that the Asp residue deemed essential for substrate binding in these last 2 enzymes has been replaced by a His in DHNA-CoA thioesterase (Fig. 2-6). On the other hand, the nucleophilic catalyst Asp of 4-HBT and gentisyl-CoA thioesterase is strictly conserved in DHNA-CoA thioesterase (Fig. 2-6). It is possible that the stringent substrate specificity of DHNA-CoA thioesterase reflects the existence, upstream in the pathway, of another aromatic CoA derivative (OSB-CoA), whose enzymatic hydrolysis would create a futile cycle and annihilate the biosynthetic flux of phylloquinone.

The slr0204 knockout is devoid of DHNA-CoA thioesterase activity and accumulates DHNA-CoA. The trace levels of phylloquinone detected in this mutant are therefore most likely attributable to the chemical decomposition of DHNA-CoA rather than to the action of moonlighting enzymes. Arguing in favor of this hypothesis, we have routinely observed that at physiological pH, pure preparations of DHNA-CoA spontaneously decomposed into DHNA and CoA. Whether enzymatic or chemical, the background hydrolysis of DHNA-CoA in vivo is not by itself sufficient to sustain adequate phylloquinone production. It results in a profound change of phylloquinone redox homeostasis and growth impairment at high light intensity. Similar phototoxicity because of the absence of phylloquinone has been reported for the *Synechocystis* mutants corresponding to DHNA phytyl transferase and DHNA synthase (the *menA* and *menB* homologs, respectively), and has been attributed to an excess reduction of photosystem II (9). The demonstration that a dedicated enzyme hydrolyses DHNA-CoA in cyanobacteria, together with the availability of a corresponding knockout mutant, opens the door to the functional characterization of the homologs of such an enzyme in green algae and flowering plants and to its identification in facultative anaerobic bacteria and archaea.

Methods

Chemicals and Reagents: DHNA-CoA synthesis was modified from Jiang et al. (19); a detailed procedure is given in *SI Materials and Methods*. OSB was prepared by dissolving 4 mg of OSB-dilactone in 550 μL 0.68 M NaOH. The solution was incubated for 3 min at 100 °C, and then neutralized with 100 μL 3.2 M HCl. OSB-dilactone was synthesized from phtalique anhydride and succinic acid in the presence of potassium acetate as described (27). Benzoyl-CoA, phenylacetyl-CoA, phylloquinone, and DHNA were from Sigma. All other chemicals were from Fisher Scientific.

Production of Recombinant Slr0204 Protein: The slr0204 gene minus its stop codon was amplified from *Synechocystis* sp. PCC 6803 genomic DNA by using the primers 5′-CATATGGGGACATTTACCTATGA-3′(forward) and 5′-CTCGAGCGGGTGATCCAAATCACAA-3′ (reverse), which contained the NdeI and XhoI restriction sites (italicized), respectively. The amplified products were ligated into pGEM-T Easy (Promega), and the construct was verified by sequencing. The 416-bp *NdeI/XhoI*-digested fragment was then subcloned into the corresponding sites of pET43.1a (Novagen) to yield an in-frame fusion upstream a 6×His-tag. This construct then was introduced in BL21 (DE3) cells. The expression and the purification of recombinant Slr0204 are detailed in *SI Materials and Methods*.

Enzyme Assays: DHNA-CoA thioesterase assays (50 μL) contained 20 mM NaH2PO4 (pH 7.0), 100 mM NaCl, 60–150 μM DHNA-CoA, and 0–15 μg proteins and were run for 5–30 min at 30°C protected from light. External standards of DHNA were incubated in the same conditions minus DHNA-CoA and enzyme. Reactions were stopped on ice with 150 μL 95% EtOH (vol/vol). The samples were centrifuged (14,000 *x g* for 10 min at 8 °C) and immediately analyzed by HPLC using the same conditions as those described for the purification of DHNA-CoA (*SI Materials and Methods*). DHNA was detected by fluorescence (Ex = 260 nm, Em = 450 nm) at a retention time of 11.6 min and was quantified relative to the external standards of DHNA. Disappearance of DHNA-CoA was monitored in parallel at A392. The hydrolysis of benzoyl-CoA, phenylacetyl-CoA, palmitoyl-CoA, and succinyl-CoA was measured spectrophotometrically with DTNB. Assays (250

μL) contained 20 mM NaH2PO4 (pH 7.0), 100 mM NaCl, 11–120 μM substrates, and 7–8 μg recombinant DHNA-CoA thioesterase. Assays with palmitoyl-CoA contained in addition 3 μM BSA. Blank samples contained no enzyme. Reactions were incubated for 30–120 min at 30 °C and were then mixed with 250 μL of an aqueous solution of 400 μM DTNB. Change in A412 compared to blank samples was measured immediately.

Generation of *slr0204* Knockout and Functional Complementation: Gene slr0204 was deleted by replacing its entire sequence with the aadA spectinomycin marker (14). For that, the flanking regions of slr0204 were amplified from *Synechocystis* sp. PCC 6803 genomic DNA by using the following pairs of primers: Pair 1, 5′-ATTGGGCGTAAATTCGCAAA-3′ (forward)/5′-GTAAACTTGCCTTTCATAGG-3′ (reverse) and Pair 2, 5′-CCTATGAAAGGCAAGTTTAC*CGTACG*GCGATTTGTGATTTGG-3′ (forward) containing a *BsiWI* site (italicized)/5′-AAGGCGTGGTGATGAAATC-3′ (reverse) for the upstream and downstream regions, respectively. The two PCR products were combined and annealed to serve as a template in a third round of amplification with forward primer of Pair 1 and reverse primer of Pair 2, and the amplified product was cloned into pGEM-T Easy (Promega). The *aadA* gene was isolated from the pMenG-Spec-KO vector (14) using *BsiWI* and was ligated in the pGEM-T Easy construct, which had been linearized with the same enzyme. Transformation of wild-type *Synechocystis* sp. PCC 6803 with this final construct and isolation of homogenous knockout clones was performed as described (28) using spectinomycin-containing medium. Replacement of slr0204 by the *aadA* marker was verified by PCR amplification of genomic DNA. For the functional complementation of the *slr0204* knockout, gene slr0204 was amplified from *Synechocystis* sp. PCC 6803 genomic DNA by using the primers 5′-GGAATTCTG*CATATG*GGGACATTTACCTATG-3′ (forward) and 5′-CG*GGATCC*TCACGGGTGATCCAAATC-3′ (reverse), which contained the *NdeI* and *BamHI* restriction sites (italicized), respectively. The amplified product was cloned into *NdeI/BglII*-digested pSynExp-2 vector, for expression under the control of the *psbA2* cyanobacterial promoter (29). This construct was introduced into the *slr0204* knockout, and transformed cells

were selected on BG-11 medium containing spectinomycin and chloramphenicol. Incorporation of slr0204 was verified by PCR on genomic DNA. For growth comparison experiments, *Synechocystis* cells were grown on solid BG-11 medium containing 5 mM glucose and no antibiotics.

Metabolite Analyses: Cells were grown at a light intensity of 80 $\mu E \cdot m^{-2} \cdot s^{-1}$ to A_{730} approximately 0.8–1.2, in liquid BG-11 medium containing 5 mM glucose with or without addition of DHNA or OSB, and with the appropriate antibiotics. Subsequent operations were at room temperature and protected from light to avoid photodegradation of naphthoquinones. The analysis of phylloquinol and phylloquinone in *Synechocystis* cells has been adapted from that described for plant tissues (30) and is detailed in *SI Materials and Methods*. DHNA-CoA in *Synechocystis* extracts was quantified after its enzymatic conversion into DHNA; a detailed procedure is given in *SI Materials and Methods*.

SI Materials and Methods

DHNA-CoA Synthesis: All steps were protected from light. DHNA (204 mg), 230 mg *N*-hydroxysuccinimide (NHS), and 384 mg 1–1-(3-dimethylaminopropyl)-3-ethylcarbodiimide hydrochloride were mixed in 5 mL tetrahydrofuran under N_2 atmosphere, and stirred for 8 h at room temperature. The mixture was then evaporated to dryness, resuspended in 2 mL dichloromethane, and applied to a silica gel 60 column (50 cm x 5 cm). DHNA-NHS was eluted from the column with approximately 70 mL ethyl acetate/hexane (5:1) and evaporated to dryness. DHNA-NHS (16 mg) was then resuspended in 1.5 mL tetrahydrofuran and mixed with 1.5 mL of an aqueous solution of 0.013 M Na_2CoA-SH (pH 8.0) under N_2 atmosphere. The mixture was stirred for 15 min at room temperature, while keeping pH approximately 8.0 with 0.2 M NaOH. Tetrahydrofuran was evaporated with a gentle stream of N2 at 4°C, and insoluble materials were removed by centrifugation before storage at -80 °C. DHNA-CoA was purified by HPLC from thawed reaction mixtures using a Zorbax Eclipse XDB-C18 column (5 μm, 4.6 x 150 mm). The column was eluted at 30°C with a linear gradient, starting from 20% methanol and 80% 50 mM

sodium acetate, pH 5.9, to 100% methanol over 30 min at a flow rate of 0.75 mL min^{-1}. Absorbance was monitored at 392 nm; DHNA-CoA eluted at 16.5 min. DHNA-CoA containing fractions were collected, evaporated with gaseous N_2, and resuspended in 0.2 mL degassed 20 mM NaH_2PO_4 (pH 7.0), 100 mM NaCl to be used immediately.

Expression and Purification of Recombinant Slr0204: Starter cultures (5 mL) grown in LB medium containing 150 g mL^{-1} ampicillin were used to inoculate 500 mL prewarmed LB medium without antibiotic. When A_{600} reached approximately 0.9, isopropyl-1-thio-β-D-galactopyranoside was added to a final concentration of 100 μM, and the incubation was continued for 6 h at 28 °C. Subsequent operations were at 4 °C. Cells were harvested by centrifugation, resuspended in 8 mL extraction buffer [50 mM NaH_2PO_4, (pH 8.0), 300 mM NaCl, 10% glycerol (vol/vol), 10 mM imidazole], and disrupted with 0.1 mm zirconia/silica beads in a MiniBeadbeater (BioSpec Products) at 5,000 rpm for 5 x 20 s. The extracts were centrifuged (14,000 _ *g* for 10 min), and the histidine-tagged protein was isolated under native conditions with Ni-NTA His-Bind resin (Novagen) following the manufacturer's recommendations. The purified protein was immediately desalted on a PD-10 column (GE Healthcare) equilibrated in 20 mM NaH_2PO_4 (pH 7.0), 100 mM NaCl, 10% glycerol (vol/vol). The desalted fraction was frozen in liquid N_2 and stored at -80 °C, which preserved the enzymatic activity.

Quantification of Phylloquinol and Phylloquinone in *Synechocystis* Extracts: For phylloquinone analysis, cells from 1.8-mL culture aliquots were harvested by centrifugation, washed once with 1mL BG-11 medium, and finally resuspended in 250 μL BG-11 medium. A 100 μL aliquot was then added to 550 μL 95%(vol/vol) ethanol containing 142 mM β-mercaptoethanol, spiked with 0.25–0.5 nmol menaquinol/menaquinone mixture as an internal standard and vortexed briefly. The lysate was cleared by centrifugation and immediately analyzed by HPLC with post-column chemical reduction coupled to fluorescence detection (Excitation: 238 nm, Emission: 426 nm). Cells were quantified by absorbance at 730 nm on a dilution of the remaining resuspended pellet and using the formula 0.25 unit A730 N108 cells.

DHNA-CoA Analysis in *Synechocystis* Extracts: Cells from 500-mL cultures were harvested by centrifugation, washed with 50 mL BG-11 medium, resuspended in 8mL 50mM NaH_2PO_4 (pH 7.0),100 mM NaCl, 5 mM DTT, and then disrupted with 0.1 mm zirconia/silica beads in a MiniBeadbeater (5,000 rpm for 5 x 20 s). The lysates were cleared by centrifugation and then split in half. The first half was incubated for 1 h at 30 °C with 21 μg recombinant Slr0204, whereas the second half was kept on ice. At the end of the incubation, the lysates were acidified to pH 4.0 with HCl, and were extracted twice with 12 mL ethyl acetate. The ethyl acetate fractions were combined and evaporated to dryness with gaseous N_2. The residue was resuspended in 250 μL 100% methanol and analyzed by HPLC-fluorescence using the same solvent system as that described for the purification of DHNA-CoA
(*SI Materials and Methods*). The level of DHNA-CoA was calculated by subtracting the amount of DHNA detected in the enzyme-treated sample (hydrolyzed DHNA-CoA + endogenous DHNA) with that of the sample kept on ice (endogenous DHNA).

Acknowledgments

We thank Dr. Patrick Dussault of the Department of Chemistry at UNL for assistance with the preparation of DHNA-CoA and Dr. Edgar Cahoon of the Department of Biochemistry at UNL for the gift of the pMenG-Spec-KO and pSynExp-2 vectors. This work was made possible by startup funds from the Nebraska Tobacco Settlement Biomedical Research Development Funds and the Center for Plant Science Innovation at the University of Nebraska-Lincoln.

References

1. Sakuragi Y, Bryant DA (2006) in *Photosystem I: The Light-Driven Plastocyanin:Ferredoxin Oxidoreductase. Advances in Photosynthesis and Respiration*, ed Golbeck JH (Springer, Dordrecht), pp 205–222.
2. Collins MD, Jones D (1981) Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implications. *Microbiol Rev* 45:316–354.
3. Wallace BJ, Young IG (1977) Role of quinones in electron transport to oxygen and nitrate in *Escherichia coli.* Studies with a *ubiA- menA_* double quinone mutant. *Biochim Biophys Acta* 461:84–100.
4. Fromme P, Grotjohann I (2006) in *Photosystem I: The Light-Driven Plastocyanin:Ferredoxin Oxidoreductase. Advances in Photosynthesis and Respiration*,

ed Golbeck JH (Springer, Dordrecht), pp 47–69.
5. Vermeer C, et al. (2004) Beyond deficiency: Potential benefits of increased intakes of vitamin K for bone and vascular health. *Eur J Nutr* 43:325–335.
6. Ichikawa T, Horie-Inoue K, Ikeda K, Blumberg B, Inoue S (2006) Steroid and xenobiotic receptor SXR mediates vitamin K2-activated transcription of extracellular matrixrelated genes and collagen accumulation in osteoblastic cells. *J Biol Chem* 281:16927–16934.
7. Hiratsuka T, et al. (2008) An alternative menaquinone biosynthetic pathway operating in microorganisms. *Science* 321:1670–1673.
8. Meganathan R (2001) Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): A perspective on enzymatic mechanisms. *Vitam Horm* 61:173–218.
9. Johnson TW, et al. (2000) Recruitment of a foreign quinone into the A(1) site of photosystem I. I. Genetic and physiological characterization of phylloquinone biosynthetic pathway mutants in *Synechocystis* sp. pcc 6803. *J Biol Chem* 275:8523–8530.
10. Wade Johnson T, et al. (2003) The *menD* and *menE* homologues code for 2-succinyl-6-hydroxyl-2,4-cyclohexadiene-1-carboxylate synthase and O-succinylbenzoic acid-CoA ligase in the phylloquinone biosynthetic pathway of *Synechocystis* sp. PCC 6803. *Biochim Biophys Acta* 1557:67–76.
11. Sakuragi Y, et al. (2002) Insertional inactivation of the menG gene, encoding 2-phytyl-1,4-naphthoquinone methyltransferase of *Synechocystis* sp. PCC 6803, results in the incorporation of 2-phytyl-1,4-naphthoquinone into the A(1) site and alteration of the equilibrium constant between A(1) and F(X) in photosystem I. *Biochemistry* 41:394–405.
12. Shimada H, et al. (2005) Inactivation and deficiency of core proteins of photosystems I and II caused by genetical phylloquinone and plastoquinone deficiency but retained lamellar structure in a T-DNA mutant of Arabidopsis. *Plant J* 41:627–637.
13. Gross J, et al. (2006) A plant locus essential for phylloquinone (vitamin K1) biosynthesis originated from a fusion of four eubacterial genes. *J Biol Chem* 281:17189–17196.
14. Lohmann A, et al. (2006) Deficiency in phylloquinone (vitamin K1) methylation affects prenyl quinone distribution, photosystem I abundance, and anthocyanin accumulation in the *Arabidopsis* AtmenG mutant. *J Biol Chem* 281:40461–40472.
15. Kim HU, van Oostende C, Basset GJ, Browse J (2008) The AAE14 gene encodes the *Arabidopsis* o-succinylbenzoyl-CoA ligase that is essential for phylloquinone synthesis and photosystem I function. *Plant J* 54:272–283.
16. Lefebvre-Legendre L, et al. (2007) Loss of phylloquinone in *Chlamydomonas* affects plastoquinone pool size and photosystem II synthesis. *J Biol Chem* 282:13250–13263.
17. Sharma V, Suvarna K, Meganathan R, Hudspeth ME (1992) Menaquinone (vitamin K2) biosynthesis: Nucleotide sequence and expression of the *menB* gene from *Escherichia coli. J Bacteriol* 174:5057–5062.
18. Truglio JJ, et al. (2003) Crystal structure of *Mycobacterium tuberculosis* MenB, a key enzyme in vitamin K2 biosynthesis. *J Biol Chem* 278:42352–42360.
19. Jiang M, et al. (2008) Identification and characterization of (1*R*,6*R*)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of *Escherichia coli. Biochemistry* 47:3426–3434.
20. Overbeek R, Fonstein M, D'Souza M, Pusch GD, Maltsev N (1999) The use of gene clusters to infer functional coupling. *Proc Natl Acad Sci USA* 96: 2896–2901
21. Dillon SC, Bateman A (2004) The Hotdog fold: Wrapping up a superfamily of thioesterases and dehydratases. *BMC Bioinf* 5:109.
22. Thoden JB, Holden HM, Zhuang Z, Dunaway-Mariano D (2002) X-ray crystallographic analyses of inhibitor and substrate complexes of wild-type and mutant 4-hydroxybenzoyl-CoA thioesterase. *J Biol Chem* 277:27468–27476.
23. Zhuang Z, Song F, Takami H, Dunaway-Mariano D (2004) The BH1999 protein of *Bacillus halodurans* C-125 is gentisyl-coenzymeAthioesterase. *J Bacteriol* 186:393–399.
24. Song F, Zhuang Z, Dunaway-Mariano D (2007) Structure-activity analysis of base and enzyme-catalyzed 4-hydroxybenzoyl coenzyme A hydrolysis. *Bioorg Chem* 35:1–10.
25. Dansette P, Azerad R (1970) A new intermediate in naphthoquinone and menaquinone biosynthesis. *Biochem Biophys Res Commun* 40:1090–1095.
26. Chang K-H, Liang P-H, Beck W, Scholten JD, Dunaway-Mariano D (1992) Isolation and

characterization of the three polypeptide components of 4-chlorobenzoate dehalogenase from *Pseudomonas* sp. strain CBS-3. *Biochemistry* 31:5605–5610.

27. RoserW(1884) About phtalylderivatives (*transl. from German*). II. *Ber Deut Chem Ges* 17:2770–2775.
28. Williams JGK (1988) Construction of specific mutations in photosystem II photosynthetic reaction center by genetic engineering methods in *Synechocystis* 6803. *Methods Enzymol* 167:766–778.
29. Sattler SE, Cahoon EB, Coughlan SJ, DellaPenna D (2003) Characterization of tocopherol cyclases from higher plants and cyanobacteria. Evolutionary implications for tocopherol synthesis and function. *Plant Physiol* 132:2184–2195.
30. Oostende C, Widhalm JR, Basset GJ (2008) Detection and quantification of vitamin K(1) quinol in leaf tissues. *Phytochemistry* 69:2457–2462.

CHAPTER 3

Plants Utilize Peroxisomal 1,4-dihydroxy-2-naphthoyl-CoA Thioesterases Not of Cyanobacterial Origin

Preface

After identifying the DHNA-CoA thioesterase in cyanobacteria, we thought that the identification of the plant orthologs would be fairly straightforward. As I presented in Chapter 2, homology-based searches with Slr0204 do in fact detect weakly-similar plastidial candidates in plants. At the time this fit with our assumption that the genes involved in phylloquinone biosynthesis encoded proteins targeted to the plastid, obtained from the ancient cyanobacterium that became the plastid. To our surprise, none of the plastidial homologs proposed in Chapter 2 turned out to be DHNA-CoA thioesterases. Extraordinarily, we now have evidence that plants recently-acquired a DHNA-CoA thioesterase through horizontal gene transfer that is highly diverged from its cyanobacterial counterpart. Here we also show that the enzyme appears to be peroxisomal, providing the first biochemical proof for an organellar split in plant phylloquinone biosynthesis. This fascinating story is presented here in Chapter 3 essentially as it appears in our manuscript to be submitted to *The Plant Journal*.

Phylloquinone (Vitamin K_1) biosynthesis in plants: two peroxisomal thioesterases of *Lactobacillales* origin participate in the hydrolysis of 1,4-dihydroxy-2-naphthoyl-CoA

Joshua R. Widhalm, Anne-Lise Ducluzeau, Nicole E. Buller, Laura J. Olsen, Christian G. Elowsky and Gilles J. C. Basset

Abstract

It is not known how plants cleave the thioester bond of 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA), a necessary step to form the naphthoquinone ring of phylloquinone (vitamin K_1). In fact, only recently has the hydrolysis of DHNA-CoA been demonstrated to be enzyme-driven in vivo, and the cognate thioesterase characterized in the cyanobacterium Synechocystis. With a few exceptions in certain prokaryotic (*Sorangium, Opitutus*) and eukaryotic (*Cyanidium, Cyanidioschyzon, Paulinella*) organisms, orthologs of DHNA-CoA thioesterase are missing outside of the cyanobacterial lineage. In this study, genomic approaches and functional complementation experiments identified two *Arabidopsis* genes encoding functional DHNA-CoA thioesterases. The deduced plant proteins display low percentages of identity with cyanobacterial DHNA-CoA thioesterases, and actually do not even share the same catalytic motif. GFP-fusion experiments demonstrated that the *Arabidopsis* proteins are targeted to peroxisomes, and subcellular fractionations of *Arabidopsis* leaves confirmed that DHNA-CoA thioesterase activity occurs in this organelle. *In vitro* assays with various aromatic and aliphatic acyl-CoA thioester substrates showed that the recombinant *Arabidopsis* enzymes preferentially hydrolyze DHNA-CoA. Cognate T-DNA knockout lines display reduced DHNA-CoA thioesterase activity and phylloquinone content, establishing *in vivo* evidence that the *Arabidopsis* enzymes are involved in phylloquinone biosynthesis. Extraordinarily, structure-based phylogenies coupled to comparative genomics demonstrate that plant DHNA-CoA thioesterases originate from horizontal gene transfer with a bacterium of the *Lactobacillales* order.

Introduction

In plants and certain species of cyanobacteria, phylloquinone (2-methyl-3-phytyl-1,4-naphtho-quinone or vitamin K_1; Fig. 3-1A) is a vital redox cofactor required for electron transfer in photosystem I and the formation of protein disulfide bonds (1-4). A closely related form called menaquinone (2-methyl-3-(all-transpolyprenyl)-1,4-naphthoquinone or vitamin K_2) is synthesized by red algae, diatoms, and most archaeal and bacterial species (5-7). In vertebrates, vitamin K is needed for blood coagulation, bone and vascular metabolism, and signaling (8). For humans in

particular, the phylloquinone of green leafy vegetables and vegetable oils, such as that of soybean, sunflower, olive, and canola, is the main contributor of dietary vitamin K (9).

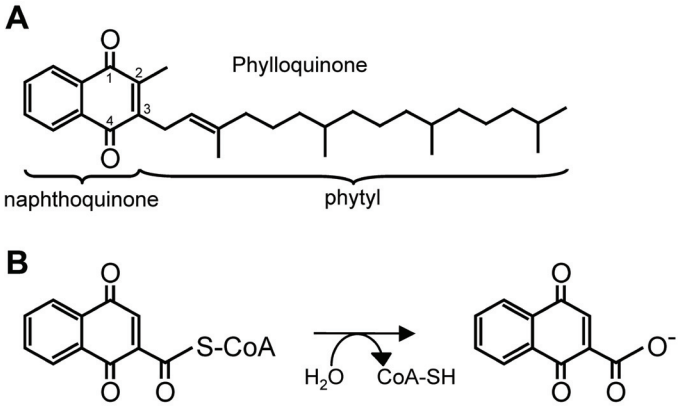


Fig. 3-1. Structure of phylloquinone and the DHNA-CoA thioesterase catalyzed reaction. (*A*), The phylloquinone molecule is bipartite, comprising a redox active naphthoquinone ring and a liposoluble phytyl side-chain. (*B*), The hydrolysis of DHNA-CoA frees the carboxyl group of DHNA for its subsequent phytylation.

Despite the importance of phylloquinone in photosynthesis and human nutrition, the molecular details of its biosynthesis in plants has only recently been explored. The immediate precursor of the redox active naphthoquinone moiety of phylloquinone is chorismate. It is first isomerized to serve as a substrate for an atypical multifunctional enzyme termed PHYLLO, that catalyzes sequential steps of addition, elimination, and aromatization suggestive of a channeling mechanism (10). PHYLLO's product, *o*-succinylbenzoate, is then activated by ligation with CoA and cyclized yielding the CoA thioester of 1,4-dihydroxy-2-naphthoate (DHNA). DHNA-CoA is subsequently hydrolyzed, and DHNA is prenylated and methylated (11-13). In agreement with radiolabelling assays showing that the prenylation and methylation reactions are associated with plastidial membranes (14-16), all the phylloquinone biosynthetic enzymes characterized so far have been shown to occur in the chloroplast (10-13, 17, 18). This apparent all-plastidial localization of the phylloquinone biosynthetic pathway has nevertheless recently been challenged by proteomic studies that identified homologs of prokaryotic DHNA-CoA synthase in *Arabidopsis* and spinach peroxisomal fractions (19, 20). GFP-reporter experiments and detection of

consensus targeting signals not only confirmed the finding, but also showed that the preceding enzyme -OSB-CoA ligase- is likely dual targeted to peroxisomes and plastids (13, 20). Although direct evidence that peroxisomal preparations actually display the aforementioned ligase and synthase activities is lacking, these preliminary results raise the intriguing possibility of a split of the phylloquinone biosynthetic pathway between chloroplasts and peroxisomes.

Besides such a fragmentary understanding of the enzymatic arrangement in plant phylloquinone biosynthesis, one step in the pathway –the hydrolysis of DHNA-CoA (Fig. 3-1B)- remains unidentified. In fact, the cleavage of the thioester bond of DHNA-CoA has long been an enigma in both prokaryotic and eukaryotic vitamin K-synthesizing organisms. Forward genetics strategies failed to isolate any mutants related to this step, and given the propensity of DHNA-CoA to spontaneously hydrolyze at physiological pH, it seemed *a priori* plausible that such a reaction was non enzymatic (21). Only recently was a dedicated DHNA-CoA thioesterase identified in the cyanobacterium *Synechocystis* PCC 6803 using a combination of phylogenomics and reverse genetics approaches (22). The enzyme, which appeared to have evolved an absolute substrate specificity for DHNA-CoA (22), was found to be related to the 4-hydroxybenzoyl-CoA thioesterase (4HBT) family of Hotdog-fold proteins (23). Relying exclusively on sequence homology to identify non-cyanobacterial DHNA-CoA thioesterase orthologs turns out to be problematic, because 4HBT-like enzymes are notorious for displaying low level of overall sequence conservation, and even dissimilar active sites, while bearing similar substrate specificity (24). Typical examples are *Pseudomonas* and *Arthrobacter* 4HBTs that catalyze the same reaction in the degradative pathway of 4-chlorobenzoate, but are classified in two separate phylogenetic subfamilies having distinct catalytic residues and quaternary structures (24-26).

In this study, we identified two *Arabidopsis* members of the 4HBT family encoding highly specific DHNA-CoA thioesterase that are targeted to peroxisomes and participate in phylloquinone biosynthesis. Using phylogenetic reconstructions, we show that these plant enzymes are not orthologous to cyanobacterial DHNA-CoA thioesterase, and likely originate from a lateral gene transfer from a bacterium of the *Lactobacillales* order.

Experimental Procedures

Chemicals and Reagents: DHNA-CoA synthesis was modified from (27). DHNA (204 mg) was activated with 230 mg of N-hydroxysuccinimide (NHS) in presence of 384 mg of 1-1-(3-dimethylaminopropyl)-3-ethylcarbo-diimide hydrochloride in 5 ml of tetrahydrofuran under N_2 atmosphere. The mixture was stirred for 8 h at room temperature, and then evaporated to dryness. The residue was resuspended in 2 mL of dichloromethane, and separated on a column (50 cm x 5 cm) of silica gel 60. DHNA-NHS was eluted with ~70 mL of ethyl acetate/hexane (5:1) and evaporated to dryness. It was then resuspended in 1.5 ml of tetrahydrofuran, and mixed with 1.5 ml of an aqueous solution of 0.013 M Na_2CoA-SH (pH 8.0) under N_2 atmosphere. The mixture was stirred for 15 min at room temperature, while keeping pH~8 with 0.2 M NaOH. Tetrahydrofuran was evaporated with a gentle stream of N_2 at 4°C, and insoluble materials were removed by centrifugation prior to storage at -80°C in 500-µL aliquots. DHNA-CoA was HPLC-purified on a Zorbax Eclipse XDB-C18 column (5 µm, 4.6 x 150 mm) immediately prior to use. The column was eluted at 30°C with a linear gradient, starting from 20% methanol and 80% of 50 mM sodium acetate pH 5.9, to 100% methanol over 30 min at a flow rate of 0.75 mL min^{-1} while monitoring A_{392}. DHNA-CoA containing fractions (~16-17 min) were collected, evaporated with gaseous N_2, and resuspended in 0.2 ml of degassed 100 mM KH_2PO_4 (pH 7.0). DHNA, benzoyl-CoA, phenylacetyl-CoA, and menaquinone-4 were from Sigma (Saint-Louis,MO). Phylloquinone was from MP biomedicals (Solon, OH). Unless otherwise mentioned, all other reagents were from Fisher Scientific.

Plant Material and Growth Conditions: *Arabidopsis* T-DNA insertion lines SAIL_1253_B02 (*At1g48320*) and SAIL_315_C08 (*At5g48950*) were obtained from the *Arabidopsis* Biological Resource Center at the Ohio State University. Seeds were allowed to germinate on Murashige-Skoog solid medium and transferred to potting mix in a growth chamber at 22°C (100 µE m^{-2} s^{-1}) in 16-h days for 6 weeks. The double knockout was obtained by crossing individual homozygous mutants. Plants were genotyped using primers LP, 5'-ATCCAATCCTCTGAAACCCTC-3'/ RP, 5'-GTGCTTACAGGAGTTGCTTCG-3' (SAIL_1253_B02) and LP, 5'-

CCATCCATTTGTATACCCGTG-3'/ RP, 5'-TGTTTTGATGCAATATCGTGTG-3' (SAIL_315_C08), and T-DNA specific primer LB2, 5'-GCTTCCTATTATATCTTCCCAAATTACCAATACA-3'. For the preparation of chloroplasts and mitochondria, *Arabidopsis* seedlings (Col-0) were grown at 22°C (100 μE m^{-2} s^{-1}) in 10-h days for 2 weeks. For the isolation of peroxisomes, *Arabidopsis* plants were grown in 16-h days for 4 weeks.

Purification of *Arabidopsis* Organelles: For the isolation of chloroplasts and mitochondria, two week-old *Arabidopsis* seedlings were de-starched in darkeness for 18h prior to tissue disruption. Chloroplasts were purified on a Percoll-gradient as described (28) except that ascorbate and bovine serum albumin were omitted from the extraction and wash buffers. For the preparation of mitochondria, 13 g of leaves were homogenized in 75 mL of homogenization buffer (20 mM sodium pyrophosphate-HCl, pH 7.5, 1 mM EDTA, 300 mM mannitol) in a standard blender (3 x 5 s pulses). All steps were at 4°C. The homogenate was filtered through one layer of miracloth, and the remaining solid material was re-extracted with 25 mL of homogenization buffer. The filtrate was centrifuged (1,500 x g, 10 min), and the supernatant was collected to be re-centrifuged (10,000 x g, 20 min). The resulting pellet was recovered and resuspended in 2 mL of sample buffer (10 mM HEPES-KOH pH 7.2, 1mM EDTA, 300 mM mannitol). The sample (2 x 1 ml) was then layered over a discontinuous density gradient consisting of 60% (1.5 mL) and 28% (6 mL) of percoll prepared in sample buffer. After centrifugation (41,000 x g, 40 min), the mitochondrial fractions (approximately 2 x 1.5 mL) were collected at the interface of the percoll layers. Peroxisomes were essentially prepared as described (29). Marker enzymes, glyceraldehyde-3-phosphate dehydrogenase (GAPDH), fumarase, and catalase were assayed as described (30).

Functional Complementation of *Synechocystis*: *Arabidopsis* cDNA clones G61320 (At1g48320), U89448 (At5g48950), G21545 (At3g61200), G12298 (At1g04290), G60738 (At2g29590), G13733 (At3g16175), G67253 (At5g48370), G67273 (At2g30720), U14083 (At1g68260), U84163 (At1g35250), and U14921 (At1g35290) were obtained from the *Arabidopsis* Biological Resource Center. No clone was available for At1g68280, and no corresponding cDNAs

were amplified from total leaf RNA. Each cDNA was PCR-amplified for subcloning between the 5'-*NdeI*/3'-*BamHI*, or for clone U89448 5'-*NdeI*/3'-*BglII*, sites of expression vector pSynExp-2 (31). A list of the corresponding primers used for the amplifications is provided (Table 3-1). All DNA constructs were verified by sequencing and were introduced into *Synechocystis slr0204::Spec* knockout (22) as described (32). Transformed cells were selected for resistance to spectinomycin and chloramphenicol. Incorporation of cDNAs was verified by PCR on genomic DNA. Plates were incubated at 22°C, 30 μE m^{-2} s^{-1}. For photosensitivity experiments, replicated plates without glucose or antibiotics were transferred at 150 μE m^{-2} s^{-1}.

Table 3-1. Primer sets used for the subcloning of *Arabidopsis* cDNA clones in expression vector pSynExp-2.*NdeI* and *BamHI* (or *BglII* for clone U89448) restriction sites in forward and reverse primers,respectively, are underlined.

cDNA	Forward Primer	Reverse Primer
At1g48320	5'-GGAATTCTGCATATGGATTCTGCATCGTCCAA-3'	5'- CGGGATCCCTACAACTTTGCGACCATTTT-3'
At5g48950	5'-GGAATTCTGCATATGGATCCAAAATCGCCGG-3'	5'- GAAGATCTTTATAATTTGGATATAACTTTCTTC-3'
At3g61200	5'-GGAATTCTGCATATGAATTGGGCAAAATTCCGT-3'	5'- CGGGATCCTTAAAGCTTAGAGATGGGAGA-3'
At1g04290	5'-GGAATTCTGCATATGGATCTAGAATCGGTGAAA-3'	5'- CGGGATCCTCAAAGGTTGGAACGAGGAG-3'
At2g29590	5'-GGAATTCTGCATATGATGGAGAAAATTATGGAGT-3'	5'- CGGGATCCTCAGAGCTTACTAGCTTGTCT-3'
At3g16175	5'-GGAATTCTGCATATGGTAGATCCAACGATCCA-3'	5'- CGGGATCCTCAAAGCTTACTAAGAGTAGAA-3'
At5g48370	5'-GGAATTCTGCATATGAATTCCCCAAGACCCAT-3'	5'- CGGGATCCTTACTGAAGAATTGTGCCCAC-3'
At2g30720	5'-GGAATTCTGCATATGAGATCTTCAGCGGGAAAG-3'	5'- CGGGATCCTCAAGGCAATGAGATGGGTC-3'
At1g68260	5'-GGAATTCTGCATATGTTTCTTCAGGTTACCGG-3'	5'-CGGGATCCTCAAACGGCGTCGTCTTGG-3'
At1g35250	5'-GGAATTCTGCATATGTTTCAAGCTACCAGCAC-3'	5'-CGGGATCCTCAATCGACCAAGTGTTGAC-3'
At1g35290	5'-GGAATTCTGCATATGCTTAAAGCTACCGGCAC-3'	5'-CGGGATCCTCAATATTCGACAACGTGTTG-3'

Subcellular localization: Full-length *At5g48950* cDNA was subcloned into pK7WGF2 (33) using Gateway™ technology, resulting in an in-frame fusion with the C-terminal end of GFP. A KAT2-eqFP611 peroxisomal marker cassette, encoding for a C-terminal fusion of *Arabidopsis* 3-ketoacyl-CoA thiolase 2 (residues 1-99) with RFP under the control of the 35S promoter (34), was subcloned into the *EcoR*I and *Pst*I sites of binary vector PZP212 (35). 35S::GFP-*AtDHNAT2* and 35S::KAT2-eqFP611 constructs were individually electroporated into *Agrobacterium tumefaciens* C58C1, for subsequent co-infiltration into the leaves of *Nicotiana benthamiana*. Mesophyll leaf

tissues were imaged by confocal microscopy two days later.

Expression and Purification of Recombinant ATDHNAT1 and 2: *At1g48320* and *At5g48950* cDNAs were subcloned minus their stop codons using Gateway™ technology in pET-DEST42 vector (Invitrogen) for C-terminal fusion with a 6xHis-tag. Constructs were introduced into *E. coli* BL21-CodonPlus (DE3)-RIL cells (Agilent). Starter cultures containing ampicillin were used to inoculate 500 ml of pre-warmed LB medium without antibiotic. When A_{600} reached ~0.9, isopropyl-1-thio-β-D-galactopyranoside (500 µM) was added, and incubation was continued for 2h at 30°C. Subsequent operations were at 4°C. Cells were harvested by centrifugation, resuspended in 8 mL of extraction buffer (50 mM NaH_2PO_4 (pH 8.0), 300 mM NaCl, 10% glycerol (vol/vol), 10 mM imidazole), and disrupted with 0.1 mm zirconia/silica beads in a MiniBeadbeater (BioSpec Products, Inc, Bartlesville, OK, USA) at 5,000 rpm for 5 x 20 s. The extracts were centrifuged (14,000 x g, 10 min) and the recombinant proteins were purified under native conditions with Ni-NTA His-Bind resin (Novagen) following the manufacturer's recommendations. Isolated proteins were immediately desalted on a PD-10 column (GE Healthcare) equilibrated in 100 mM KH_2PO_4 (pH 7.0), 10% glycerol (vol/vol). Desalted fractions were frozen in liquid N_2 and stored at –80°C.

Preparation of *Arabidopsis* crude protein extracts: *Arabidopsis* leaves (0.75 g) were flash-frozen in liquid nitrogen and ground to a fine powder with pestle and mortar. The powder was transferred to a 40 mL screw-cap tube and thawed with 2.25 mL of 100 mM KH_2PO_4 pH 8.0, 5% (w/vol) polyvinylpolypyrrolidone, 5 mM freshly-prepared DTT. Samples were centrifuged (10,000 x g; 10 min at 4°C) to pellet debris. Supernatants were recovered and desalted on a PD-10 column equilibrated in 100 mM KH_2PO_4 (pH 7.0), 5 mM DTT, 10% glycerol (vol/vol).

Enzyme assays: DHNA-CoA thioesterase assays (50 µL) contained 100 mM KH_2PO4 (pH 7.0), 5 mM DTT, 35-90 µM DHNA-CoA, and 0-49 µg of proteins, and were incubated in the dark for 10-20 min at 30°C. Negative controls containing boiled proteins, and external standards of DHNA

were incubated in parallel. Assays were terminated with the addition of 150 µL ice-cold 95% ethanol (vol/vol). Samples were then centrifuged (16,000 x g; 5 min, 8°C) and immediately analyzed by HPLC with fluorescence and diode array detection modules as previously described (22). The hydrolysis of the benzoyl-CoA, phenylacetyl-CoA, palmitoyl-CoA and succinyl-CoA substrates was measured spectrophotometrically using the DTNB-derivation method. Assays (375 µL) contained 100 mM KH_2PO_4 (pH 7.0), 90 µM substrates, and 0.65-2.7 µg of recombinant AtDHNAT1 and 2. Assays with palmitoyl-CoA contained in addition 3 µM of BSA. Blank samples containing no enzyme or no substrate were included. Reactions were incubated for 60- 180 min at 30°C, and were then mixed with 375 µL of an aqueous solution of 400 µ M DTNB. Changes in A_{412} compared to blank samples were read after 5 minute incubation.

Phylloquinone Analyses: All steps were carried out in dimmed light to avoid photodegradation of naphthoquinone species. *Synechocystis* cells (1.8 ml) were harvested by centrifugation, and resuspended in 225 µL of BG-11 medium (32). Cells were quantified by absorbance at 730 nm using the formula 0.25 unit A_{730} = 10^8 cells. A 150-µL aliquot was then added to 700 µL of 95% (vol/vol) ethanol and 220 µL of water into a pyrex screw-cap tube, spiked with 75 pmoles of menaquinone-4 as an internal standard. *Arabidopsis* leaves (8-21 mg of fresh weight) were spiked with 150 pmoles of menaquinone-4 as internal standard, and were homogenized in 1 mL of 100% (vol/vol) methanol using a pyrex tissue grinder. The extract was then transferred to a pyrex tube containing 0.6 mL of water. *Synechocystis* and *Arabidopsis* extracts were then partitioned with 5 mL of hexane. Upper phases were transferred into a new tube, and thenevaporated to dryness under a gentle stream of gaseous N_2. The residue was redissolved into 200 µL of ethanol. The samples (100 µl) were analyzed by HPLC on a 5 µm Discovery C-18 column (250 x 4.6 mm, Supelco) thermostated at 30°C and eluted in isocratic mode at a flow rate of 1 mL min^{-1} with methanol: ethanol (80:20, vol/vol) containing 1 mM sodium acetate, 2 mM acetic acid, 2 mM $ZnCl_2$. Naphthoquinone species were detected fluorometrically (238 nm and 426 nm for excitation and emission, respectively) after reduction into a post-column chemical reactor (70 x 1.5 mm) packed with -100 mesh zinc powder (Aldrich). Phylloquinone and

menaquinone-4 were quantified according to external calibration standards, and data were corrected for recovery of the internal standard.

Results

***Arabidopsis* genes *At1g48320* and *At5g48950* Encode for Members of the 4HBT Family That Fully Complement *Synechocystis* DHNA-CoA Thioesterase Knockout:** Searching the Pfam database of protein families (36) for entries containing a predicted 4HBT domain identified 12 hotdog-fold proteins in *Arabidopsis*. Cognate fulllength cDNAs were obtained for all of them, except the putative product of gene *At1g68280* (see experimental procedures). Mining expressed sequence tag and microarray databases did not detect any hits for *At1g68280*. Analysis of the genomic context of *At1g68280* showed that this gene actually occurs as a tandem repeat of gene *At1g68260* (the corresponding deduced proteins being 83% identical) indicative of a recent event of gene duplication. Because such features typify *At1g68280* as a pseudogene, it was not investigated further. The other 11 cDNAs were individually sub-cloned into expression vector pSynExp-2 under the control of the cyanobacterial *psbA2* promoter (31) and introduced into *Synechocystis* strain Δ*slr0204*. This strain is a DHNA-CoA thioesterase knockout that lacks phylloquinone and is photosensitive (22). Out of the 11 aforementioned cDNAs, two – corresponding to genes *At1g48320* and*At5g48950*– restored cell growth at high-light intensities, as did reintroduction of gene *slr0204*, providing initial genetic evidence for the existence of functional plant DHNA-CoA thioesterases (Fig. 3-2A). The deduced At1g48320 and At5g48950 proteins share 63% identity and 78% similarity indicating that they are likely paralogs. Remarkably, each of these *Arabidopsis* proteins shows low percentages of homology (~15% identity/ ~30% similarity) with *Synechocystis* Slr0204. No complementation was seen with any of the other cDNAs or the vector alone control, and all clones grew at low light intensity (Fig. 3-2A). To confirm these results, the phylloquinone content of each clone was quantified by HPLC. Only cells expressing the At1g48320, At5g48950 and endogenous Slr0204 proteins displayed phylloquinone levels similar to that of the wild type reference strain (Fig. 3-2B). The At1g48320

and At5g48950 proteins were named *Arabidopsis thaliana* DHNA-CoA thioesterase 1 and 2, respectively (abbreviated hereafter as AtDHNAT1 and AtDHNAT2).

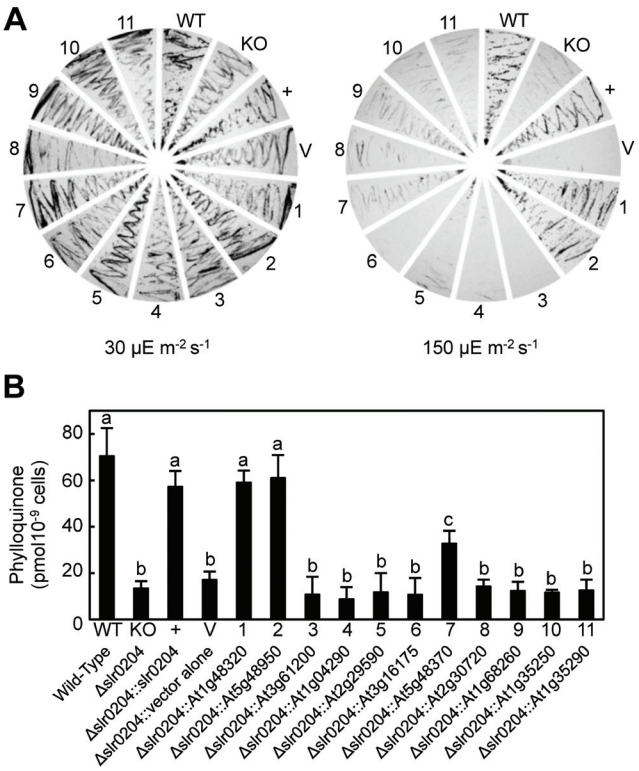


Fig. 3-2. Identification of two *Arabidopsis* 4-HBT-like proteins that encode functional DHNA-CoA thioesterases. (*A*), Functional complementation of *Synechocystis* DHNA-CoA thioesterase knockout (Δ*slr0204*). Similar number of cells of wild-type (WT), mutant (KO), and mutant transformed with vector alone (V) or native gene *Slr0204* (+) or *Arabidopsis* cDNAs encoding predicted 4-HBT domain containing proteins (clones 1-11) were plated on BG-11 without glucose or antibiotics. Plates were incubated at 22°C for 2 weeks under low (30 µE m-2 s-1) or high (150 µE m-2 s-1) light intensities. (*B*), HPLC-fluorescence quantification of phylloquinone in *Synechocystis* extracts. Clone nomenclature is the same as in panel A. Data are means ± SE of three replicates. Differing letters above columns indicate statistically different means determined from Fisher's least significant difference (LSD) test ($p < \alpha = 0.05$) from an analysis of variance (ANOVA).

AtDHNAT1 and AtDHNAT2 Occur in Peroxisomes, so Does DHNA-CoA Thioesterase Activity: AtDHNAT1 and AtDHNAT2 both have C-terminal tripeptides (AKL and SKL, respectively) typifying peroxisomal targeting signals type 1 (37), and are *de facto* listed in the AraPerox database of putative proteins of *Arabidopsis* peroxisomes (http://www3.uis.no/araperoxv1/; ref 38). Similarly, a survey of the TAIR database (http://www.*Arabidopsis*.org/index.jsp) confirmed that signature fragments of AtDHNAT1 have

been identified through large-scale proteomics experiments in purified peroxisomes, and that the N-terminally eYFP-tagged protein is targeted to this organelle (29). To verify the subcellular localization of AtDHNAT2, its full-length cDNA was fused to the C-terminal end of GFP. Coexpression of this fusion protein with an RFP-tagged peroxisomal marker resulted in a distinctive punctate pattern and co-localization of the green and red pseudocolors (Fig. 3-3A, B, D). No GFP-associated fluorescence was observed in plastids (Fig. 3-3C, D).

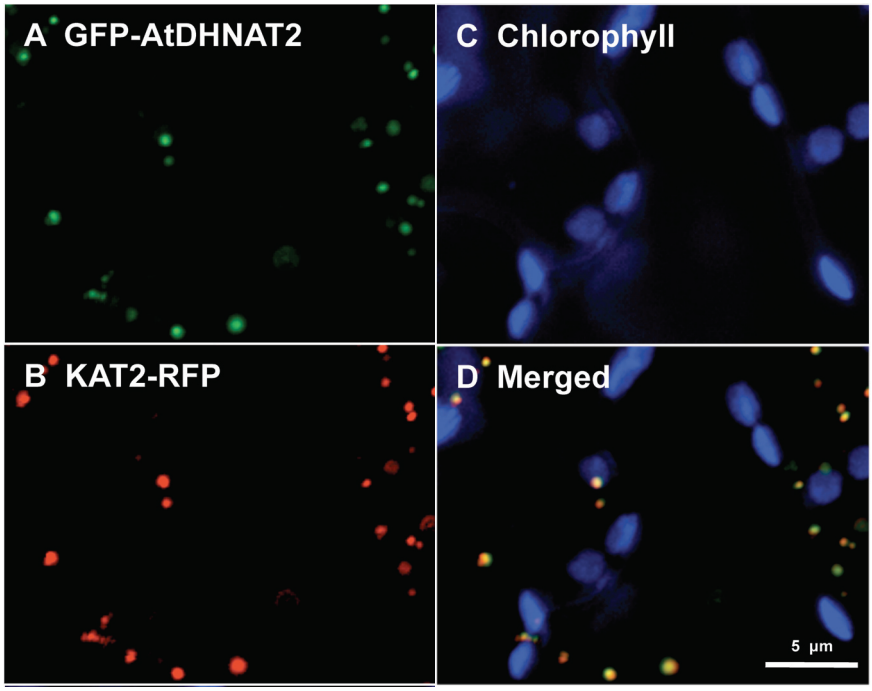


Fig. 3-3. Subcellular localization of AtDHNAT2. (*A*), green pseudocolor of GFP-tagged AtDHNAT2; (*B*), red pseudocolor of peroxisomal marker RFP-tagged 3-keto-acyl-CoA thiolase 2 (KAT2; fragment 1-99); (*C*), blue pseudocolor of chlorophyll. (*D*), overlay.

To further investigate the subcellular distribution of plant DHNA-CoA thioesterase activity, chloroplasts, mitochondria, and peroxisomes were isolated from *Arabidopsis* leaves, using assays of marker enzymes to monitor the integrity and enrichment of each organelle preparation (Table 1). Of the three purified organelles, DHNA-CoA thioesterase activity was detected only in peroxisomes, showing similar enrichment and recovery as catalase (Table 3-2). These data indicate that *Arabidopsis* DHNA-CoA thioesterase activity co-purifies exclusively with peroxisomes coinciding with the subcellular localization of AtDHNAT1 and AtDHNAT2.

Table 3-2. *Arabidopsis* DHNA-CoA thioesterase activity co-purifies with peroxisomes. DHNA-CoA thioesterase and marker enzyme activities were assayed in crude extracts (CE), stromal fraction of percoll-purified chloroplasts (CP), matrix fraction of percoll-purified mitochondria (MT), and matrix fraction of percoll-purified peroxisomes (PX) of *Arabidopsis* leaves. NADP-linked glyceraldehyde-3-phosphate dehydrogenase (GAPDH), fumarase, and catalase were used as marker enzymes for chloroplasts, mitochondria, and peroxisomes, respectively. Data are means of three replicate ± SE, except for the marker assays on peroxisomes, for which single measurements were performed.

	μmol min^{-1} mg^{-1} protein			nmol h^{-1} mg^{-1} protein
	GAPDH	Fumarase	Catalase	DHNA-CoA Thioesterase
CE	0.93 ± 0.1	0.41 ± 0.03	55 ± 15	2.4 ± 0.52
CP	1.6 ± 0.23	< 0.01	9.8 ± 1.4	< 0.5
MT	< 0.001	54 ± 18.9	24 ± 4.8	< 0.5
PX	< 0.001	0.11	287	7.3 ± 0.3

AtDHNAT1 and AtDHNAT2 Display Marked Substrate Preference for DHNA-CoA *in Vitro*: To study the substrate specificity of AtDHNAT1 and AtDHNAT2, 6xHis-tagged versions were expressed in *E. coli*, purified by affinity chromatography, and assayed with various acyl-CoA thioester substrates, all at the same concentration (Fig. 3-4). The specific activity of each recombinant enzyme towards DHNA-CoA was comparable to that measured for *Synechocystis* Slr0204 DHNA-CoA thioesterase assayed at a similar substrate concentration (74-118 μmol h^{-1} mg^{-1} for AtDHNAT1 and AtDHNAT2 at 90 μM DHNA-CoA vs. 102 μmol h^{-1} mg^{-1} for Slr0204 at 65 μM DHNA-CoA; Fig. 3-4; ref 22). AtDHNAT2 also displayed activity towards benzoyl-CoA, although the measured value was approximately an order of magnitude lower than that obtained with DHNA-CoA as substrate (Fig. 3-4). All the other activities measured with the aromatic acyl-CoA thioesters and the short-chain aliphatic acyl-CoA thioester succinyl-CoA were detected at trace levels (Fig. 3-4). No activity was detected against the longchain aliphatic acyl-CoA thioester palmitoyl-CoA (Fig. 3-4). AtDHNAT1 and AtDHNAT2 thus appear to have marked substrate preference for DHNA-CoA, and in that regard resemble cyanobacterial DHNA-CoA thioesterase (22).

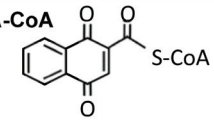
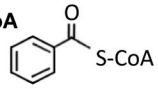
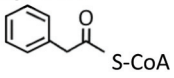
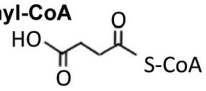
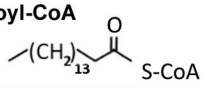
Substrate [90 µM]	Specific Activity (µmol h^{-1} mg^{-1} protein) AtDHNAT1	AtDHNAT2
DHNA-CoA 	74 ± 21	118 ± 23
Benzoyl-CoA 	1.2 ± 0.48	15.1 ± 4.7
Phenylacetyl-CoA 	0.48 ± 0.2	0.53 ± 0.2
Succinyl-CoA 	2.1 ± 0.82	2.3 ± 0.17
Palmitoyl-CoA 	< 0.001	< 0.001

Fig. 3-4. Substrate specificity of AtDHNAT1 and AtDHNAT2. Purified recombinant AtDHNAT1 and AtDHNAT2 (0.013-2.7 µg) were assayed with various aromatic and aliphatic acyl CoA-thioester substrates. DHNA-CoA thioesterase activity was monitored by direct quantification of DHNA using HPLC with diode array and fluorescence detection; the hydrolysis of the other substrates was measured spectrophotometrically by derivatization of free CoA-SH with the thiol-reagent DTNB. Data are means ± SE of three replicates.

AtDHNAT1 and AtDHNAT2 are Involved in Phylloquinone Biosynthesis in *Arabidopsis*: Two T-DNA lines from the SAIL collection (39) corresponding to insertions located in the first exon of *AtDHNAT1* (SAIL_1253_B02) and the second exon of *AtDHNAT2* (SAIL_315_C08) were identified, and confirmed by DNA genotyping (Fig. 3-5A,B). Neither of the corresponding homozygous mutants (*atdhnat1/AtDHNAT2*; *AtDHNAT1/atdhnat2*), nor the double homozygous knockout (*atdhnat1/atdhnat2*) displayed any significant phenotypic differences with wild-type plants (Fig. 3-5C). All homozygous knockout lines, on the other hand, displayed marked reduction in DHNA-CoA thioesterase specific activity compared to wild-type controls; the decrease ranging from approximately 60% for *atdhnat1/AtDHNAT2* and *atdhnat1/atdhnat2* to 25% for *AtDHNAT1/atdhnat2* (Fig. 3-6*A*). To determine the approximate contributions of AtDHNAT1 and AtDHNAT2 in the biosynthesis of phylloquinone, its content was determined in the leaves of each individual homozygous knockout, the double homozygous knockout and the wild-type control.

Phylloquinone contents were 65% of wild-type level in *atdhnat1/AtDHNAT2* and *atdhnat1/atdhnat2* plants, and 80% of wild-type level in *AtDHNAT1/atdhnat2* plants (Fig. 3-6*B*). The differences in phylloquinone content between knockout and wild-type plants were notably higher when plants were grown *in vitro*; levels reaching only 45, 52 and 30% that of wild-type in *atdhnat1/AtDHNAT2*, *AtDHNAT1/atdhnat2* and *atdhnat1/atdhnat2*, respectively (Fig. 3-6*C*).

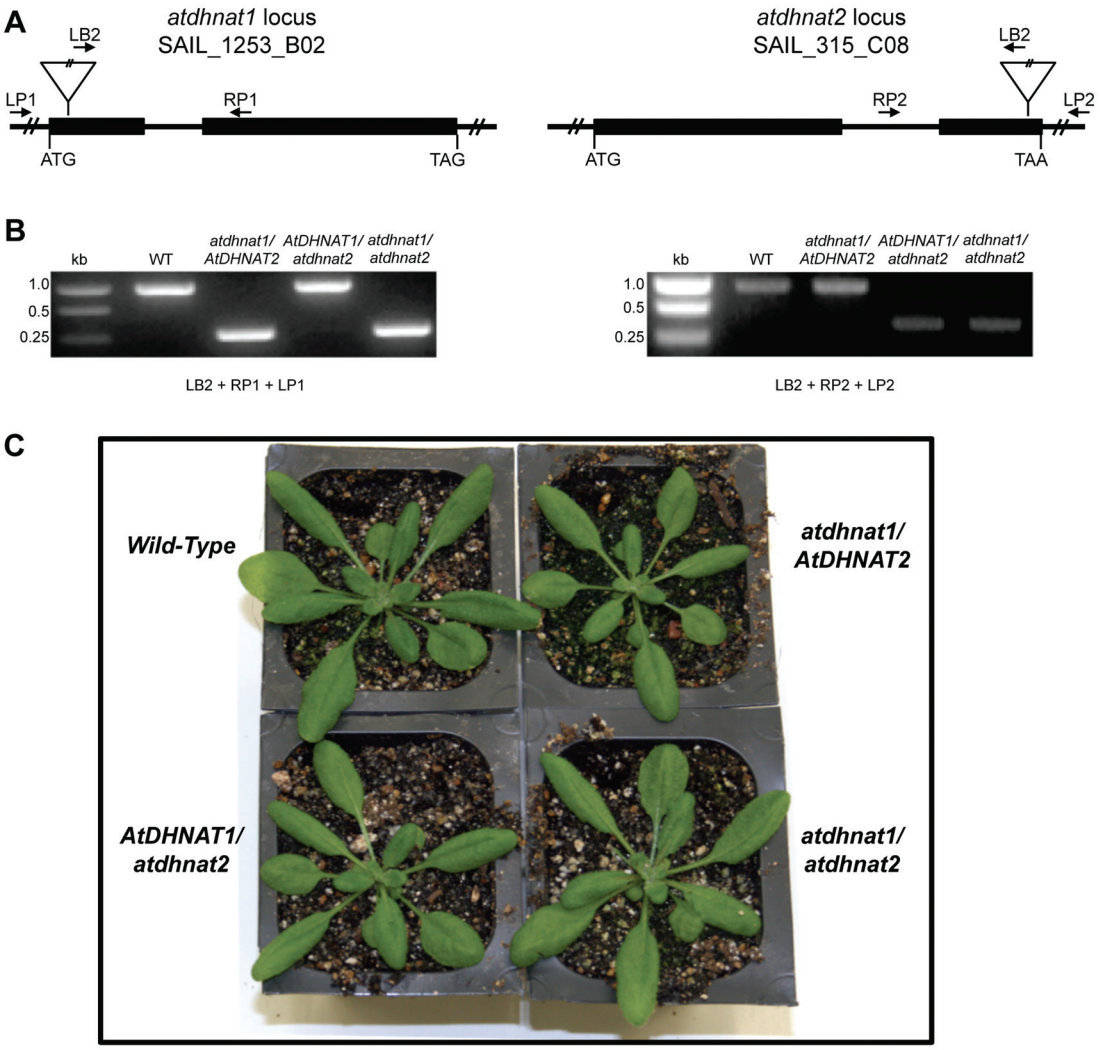


Fig. 3-5. Structure, genotyping, and phenotype associated with *atdhnat1* and *atdhnat2* loci. (*A***),** Locations of T-DNA insertions in SAIL lines 1253_B02 and 315_C08, and their corresponding genotyping primers. Black boxes and lines represent exons and introns, respectively. (*B)*, Genomic DNA-PCR using combinations of LB2, RP, and LP primers. For each locus, the upper ~1kb product corresponds to the amplification of the WT allele (RP + LP); T-DNA insertions give the lower products (RP + LB2). (*C*), Wild-type, *atdhnat1*/ *AtDHNAT2*, *AtDHNAT1/atdhnat2*, and *atdhnat1/atdhnat2 Arabidopsis* plants.

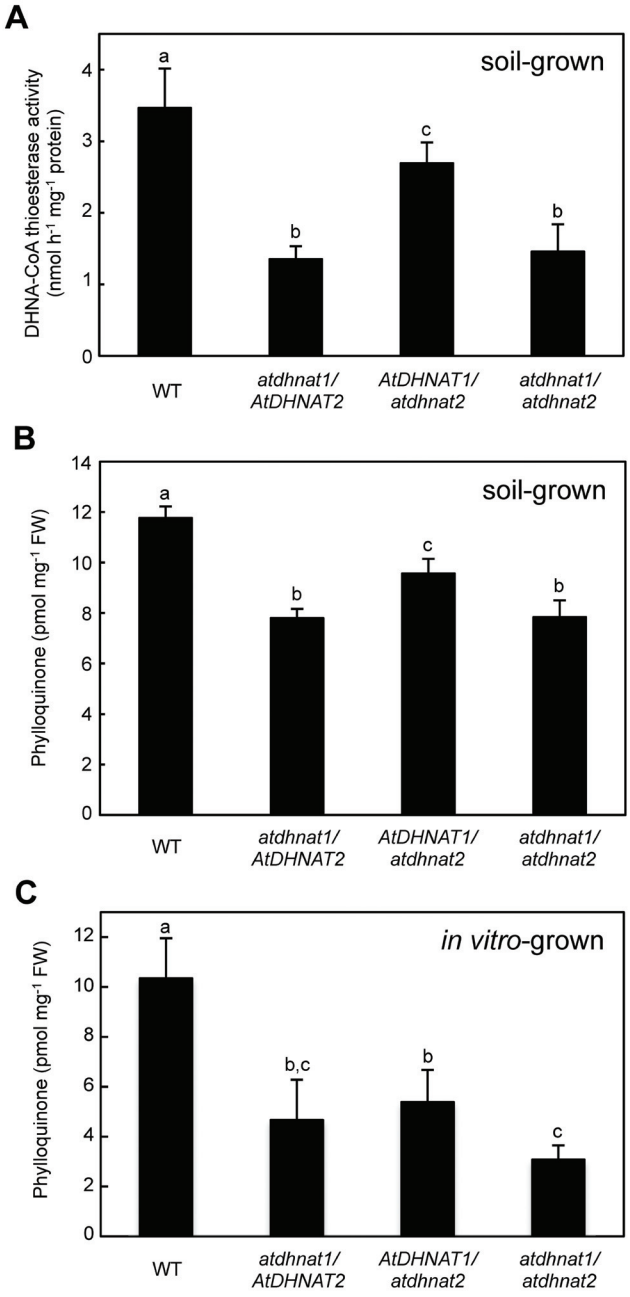


Fig. 3-6. *AtDHNAT1* and *AtDHNAT2* knockouts display reduced DHNA-CoA thioesterase activity and phylloquinone content. (*A*), Desalted extracts of soil-grown wild-type *Arabidopsis* plants (WT) and TDNA insertion lines corresponding to *AtDHNAT1* (*atdhnat1/AtDHNAT2*) and *AtDHNAT2* (*AtDHNAT1/atdhnat2*) single knockouts and to the double knockout (*atdhnat1/atdhnat2*) were assayed for DHNA-CoA thioesterase activity for 20 min at 30°C. Assays contained 85 µM DHNA-CoA and 21-49 µg of proteins. (*B*), Phylloquinone levels in the leaves of wild-type and knockout *Arabidopsis* plants grown on soil. (*C*), Same as in panel *B* but plants were grown on Murashige and Skoog medium containing 3% sucrose. Data are means ± SE of three replicates. Differing letters above columns indicate statistically different means determined from Fisher's LSD test ($p < \alpha = 0.05$) from an ANOVA.

Plant and Cyanobacterial DHNA-CoA Thioesterases Belong to Separate Phylogenetic Subfamilies: Because nothing is known about the evolutionary relationships of plant and cyanobacterial DHNA-CoA thioesterases, we reconstructed their phylogeny within families of related hotdog-fold CoA thioesterases. The bias in protein alignment due to weak conservation of amino acid sequences in this class of enzymes (24) was corrected by superposing secondary structures and catalytic sites with those of hotdog-fold CoA thioesterases of solved crystal

structures (Fig. 3-7). These reference enzymes were selected as those for which experimental evidence for DHNA-CoA thioesterase activity exists at least *in vitro* (*E. coli* YbdB; pdb file 1VH9; ref 27) or for which DHNA-CoA thioesterase function could be robustly inferred from phylogenomics evidence (*P. marinus* Slr0204; pdb file 2HX5, ref 22). The alignment was then populated with the closest homologues from various prokaryotic and eukaryotic organisms. A structure of *E. coli* Phenylacetic acid protein I (PaaI; pdb file 2FS2) was also included, for AtDHNAT1 and 2 are classified in the NCBI and TAIR databases as PaaI enzymes. The cognate tree establishes that plant and cyanobacterial DHNA-CoA thioesterases belong to strikingly different subfamilies (Fig. 3-8A), the catalytic motif of which are even unrelated (Fig. 3-7). Notably, three eukaryotic members from rhodophytes (red algae) and euglyphidae cluster within the cyanobacterial subfamily (Fig. 3-8A). These eukaryotic enzymes are plastid-encoded, and as occurs in present-day cyanobacteria have genes organized in vitamin K biosynthetic clusters, pointing to a likely origin as remnants of cyanobacterial endosymbionts (Fig. 3-8B). In stark contrast, AtDHNAT1 and 2 and their homologues in monocots, gymnosperms, mosses and Lycopodiophytes regroup monophyletically with eubacterial proteins of the *Lactobacillales* order (Fig. 3-8A). As remarkable is the finding that the corresponding genes in multiple *Lactobacillales* and other firmicutes species occur in clusters of menaquinone biosynthetic genes, thus providing unequivocal evidence that these hotdog-fold proteins encode DHNA-CoA thioesterases (Fig. 3-8B). Using the nomenclature proposed in the most recent classification of thioester-active enzymes (40), plant DHNA-CoA thioesterases appear therefore as members of the TE11/4HBT-II-type subfamily of CoA thioesterases, while their cyanobacterial counterparts represent a subfamily of their own (TE12/Slr0204-type). It is also evident that the prevailing annotation of plant DHNA-CoA thioesterases as part of the PaaI/TE13 subfamily is erroneous.

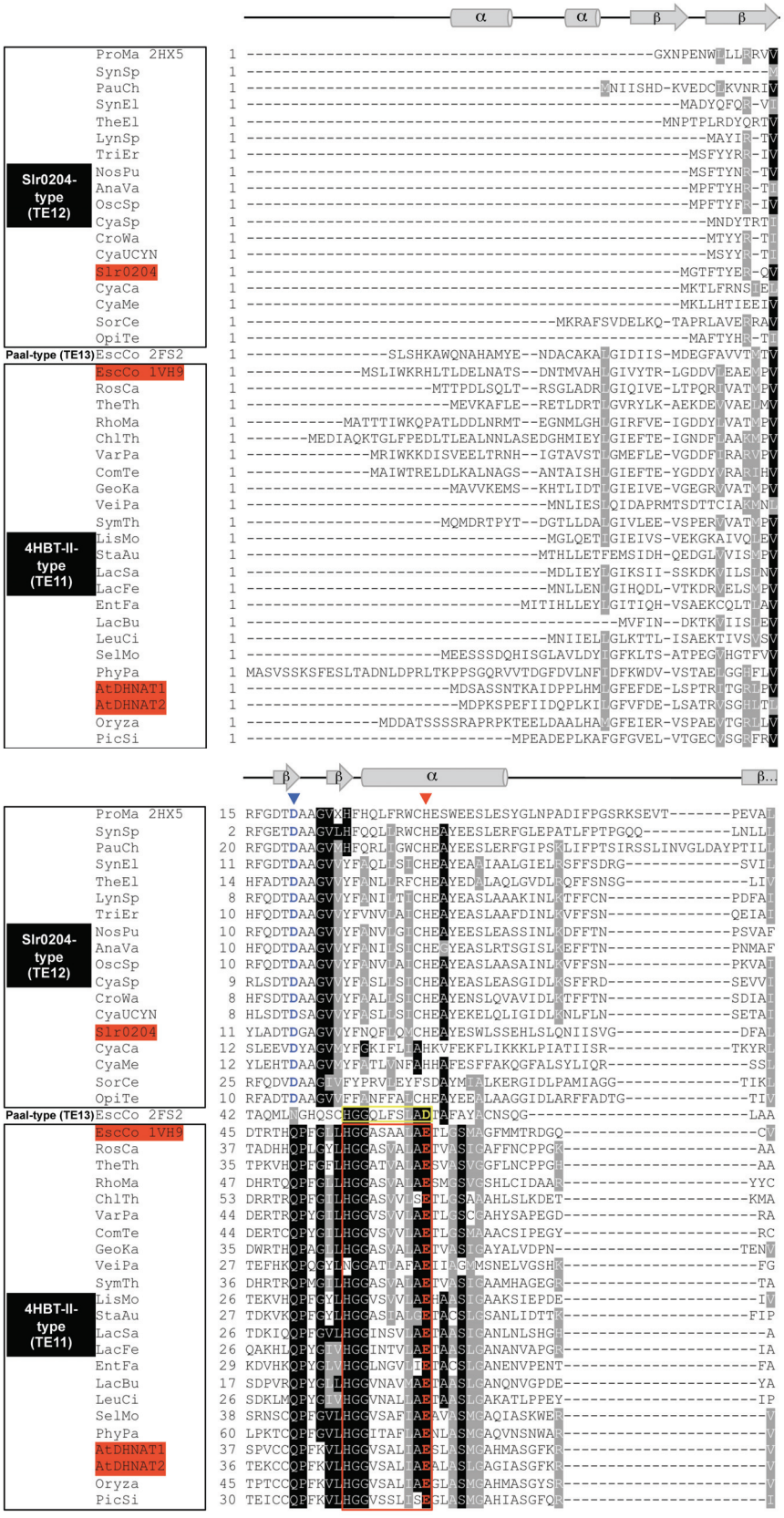
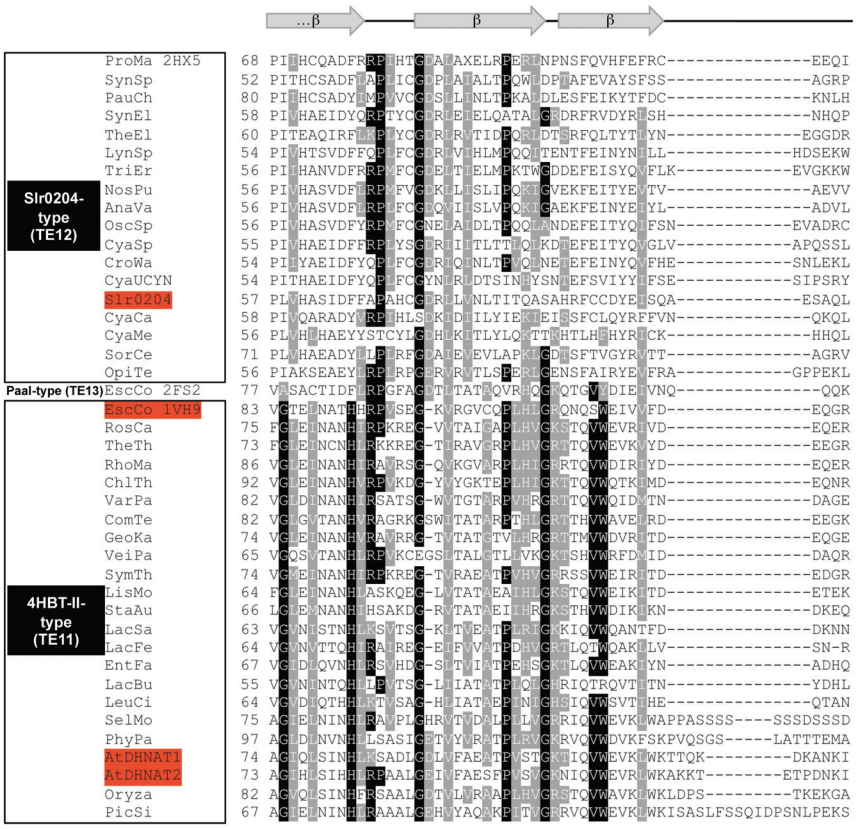


Fig. 3-7. Sequence alignment of prokaryotic and eukaryotic DHNA-CoA thioesterases and related proteins. Proteins highlighted in red have been shown to bear DHNA-CoA thioesterase activity *in vitro* and *in vivo* (Slr0204; DHNAT1-2) or *in vitro* (EscCo 1VH9). 2HX5, 2FS2, and 1VH9 indicate PDB numbers of reference tri-dimensional structures used to adjust the alignment. Corresponding secondary structures are shown on top. Identical residues are shaded in black, similar ones in grey. The TE (Thioester-active Enzymes) nomenclature and classification is from (37). The blue and red arrows point to the conserved catalytic residue in the Slr0204-type (TE12) and 4HBT-II (TE11) subfamilies, respectively. The red and yellow boxes indicate the 4HBT-II and Paal motifs, respectively, as defined in (23). Full names and taxonomic origin of species, and protein accession numbers are listed in Table 3-3.


...β
β
β
Slr0204-type (TE12)
Paal-type (TE13)
4HBT-II-type (TE11)
ProMa 2HX5
SynSp
PauCh
SynEl
TheEl
LynSp
TriEr
NosPu
AnaVa
OscSp
CyaSp
CroWa
CyaUCYN
Slr0204
CyaCa
CyaMe
SorCe
OpiTe
EscCo 2FS2
EscCo 1VH9
RosCa
TheTh
RhoMa
ChlTh
VarPa
ComTe
GeoKa
VeiPa
SymTh
LisMo
StaAu
LacSa
LacFe
EntFa
LacBu
LeuCi
SelMo
PhyPa
AtDHNAT1
AtDHNAT2
Oryza
PicSi

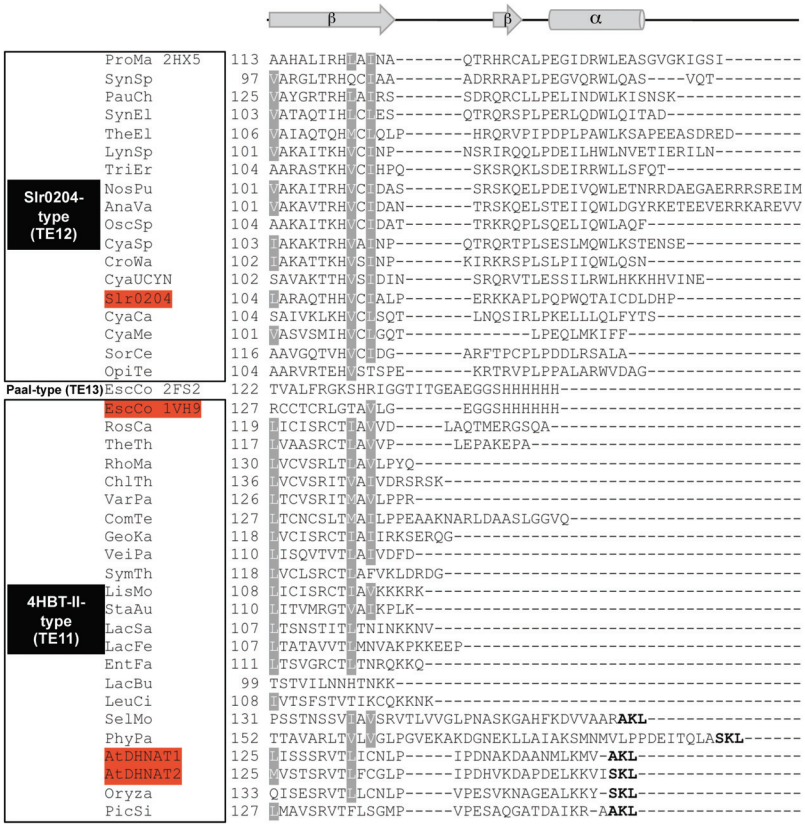

β
β
α
Slr0204-type (TE12)
Paal-type (TE13)
4HBT-II-type (TE11)
ProMa 2HX5
SynSp
PauCh
SynEl
TheEl
LynSp
TriEr
NosPu
AnaVa
OscSp
CyaSp
CroWa
CyaUCYN
Slr0204
CyaCa
CyaMe
SorCe
OpiTe
EscCo 2FS2
EscCo 1VH9
RosCa
TheTh
RhoMa
ChlTh
VarPa
ComTe
GeoKa
VeiPa
SymTh
LisMo
StaAu
LacSa
LacFe
EntFa
LacBu
LeuCi
SelMo
PhyPa
AtDHNAT1
AtDHNAT2
Oryza
PicSi

Table 3-3. Accession numbers and taxonomic origin of proteins used in phylogenetic reconstruction.

Accession number (GI)	Species	Taxonomy	Thioesterase subfamily
11465506	*Cyanidium caldarium (CyaCa)*	Rhodophyta	**Slr0204-type (TE12)**
30468143	*Cyanidioschyzon merolae strain 10D (CyaMe)*		
194476806	*Paulinella chromatophora (PauCh)*	Euglyphida	
75906622	*Anabaena variabilis ATCC 29413 (AnaVa)*	Cyanobacteria	
67924673	*Crocosphaera watsonii WH 8501 (CroWa)*		
218246677	*Cyanothece sp. PCC 8801 (CyaSp)*		
284929268	*Cyanobacterium UCYN-A (CyaUCYN)*		
119487534	*Lyngbya sp. PCC 8106 (LynSp)*		
186686472	*Nostoc punctiforme PCC 73102 (NosPu)*		
300864056	*Oscillatoria sp. PCC 6506 (OscSp)*		
114794605	*Prochlorococcus marinus Mit9313 (ProMa)* **[2HX5]**		
1001691	*Synechocystis sp. PCC 6803* ***(Slr0204)***		
56752356	*Synechococcus elongatus PCC 6301 (SynEl)*		
148243355	*Synechococcus sp. RCC307 (SynSp)*		
22298031	*Thermosynechococcus elongatus BP-1 (TheEl)*		
113477488	*Trichodesmium erythraeum IMS101 (TriEr)*		
162451288	*Sorangium cellulosum 'So ce 56' (SorCe)*	δ-proteobacteria	
182414623	*Opitutus terrae PB90-1 (OpiTe)*	Verrucomicrobia	
90109612	*Escherichia coli K-12 (EscCo)* **[2FS2]**	γ-proteobacteria	**PaaI-type (TE13)**
299533437	*Comamonas testosteroni S44 (ComTe)*	β-proteobacteria	**4HBT-II-type (TE11)**
239813088	*Variovorax paradoxus S110 (VarPa)*		
40889912	*Escherichia coli (EscCo)* **[1VH9]**	γ-proteobacteria	
268317024	*Rhodothermus marinus DSM 4252 (RhoMa)*	Bacteroidetes	
193215950	*Chloroherpeton thalassium ATCC 35110 (ChlTh)*	Chlorobi	
156233524	*Roseiflexus castenholzii DSM 13941 (RosCa)*	Chloroflexi	
46199482	*Thermus thermophilus HB27 (TheTh)*	Deinococcus-Thermus	
56379941	*Geobacillus kaustophilus HTA426 (GeoKa)*	Other Firmicutes	
217963510	*Listeria monocytogenes HCC23 (LisMo)*		
15923934	*Staphylococcus aureus subsp. aureus Mu50 (StaAu)*		
51892109	*Symbiobacterium thermophilum IAM 14863 (SymTh)*		
269797443	*Veillonella parvula DSM 2008 (VeiPa)*		
255973510	*Enterococcus faecalis T2 (EntFa)*	Lactobacillales (Firmicutes)	
227511388	*Lactobacillus buchneri ATCC 11577 (LacBu)*		
260662032	*Lactobacillus fermentum 28-3-CHN (LacFe)*		
227892451	*Lactobacillus salivarius ATCC 11741 (LacSa)*		
170018044	*Leuconostoc citreum KM20 (LeuCi)*		
33589776	*Arabidopsis thaliana* **(AtHDCT1; At1g48320)**	Plants	
149944285	*Arabidopsis thaliana* **(AtHDCT2; At5g48950)**		
115454707	*Oryza sativa Japonica Group (Oryza)*		
116786436	*Picea sitchensis (PicSi)*		
168022626	*Physcomitrella patens subsp. patens (PhyPa)*		
302804672	*Selaginella moellendorffii (SelMo)*		

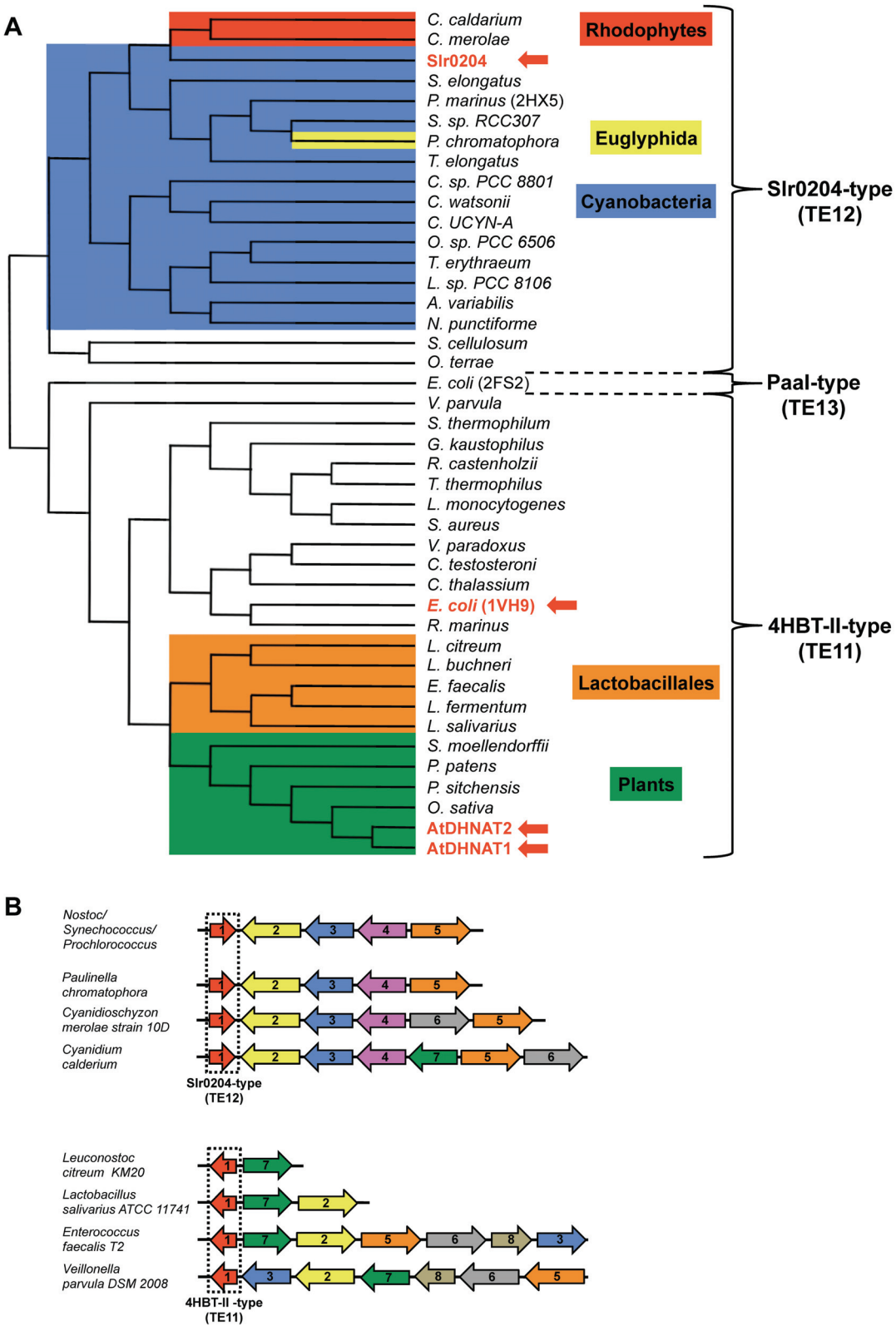


Fig. 3-8. Land plant DHNA-CoA thioesterases are not of cyanobacterial descent. (*A*), Non-exhaustive reconstruction of the structure-based maximum likelihood phylogeny of Slr0204-type, Paal-type, and 4HBT-II-type Hotdog-fold thioesterases using the MABL website (http://www.phylogeny.fr/). PDB numbers of

reference structures are given in brackets and alignments are provided as Fig. S3. Full names and taxonomic origin of species, and protein accession numbers are listed in Table S2. Red arrows point to enzymes for which there is experimental evidence of DHNA-CoA activity. (*B*), Vitamin K biosynthetic gene clusters mined from the NCBI genomic database and SEED resources for comparative genomics (http://theseed.uchicago.edu/FIG/index.cgi). Cyanobacterial, red algae, and euglyphida gene clusters are modified from ref 22. Matching colors and numbers indicate orthology. 1, DHNA-CoA thioesterase; 2, OSB-CoA ligase; 3, OSB synthase; 4, DHNA prenyltransferase; 5, isochorismate synthase; 6, SHCHC synthase; 7, DHNA-CoA synthase; 8, SEPHCHC synthase.

Discussion

We have identified two *Arabidopsis* cDNAs specifying peroxisomal DHNA-CoA thioesterases involved in the biosynthesis of phylloquinone. It is noteworthy that neither of them correspond to the four plastid-targeted proteins that we had previously proposed, based on homology searches with the cyanobacterial enzyme (Slr0204), as putative *Arabidopsis* DHNA-CoA thioesterases (22). Of those, three -At1g68260, At1g35250, At1g35290- fail to complement the *Synechocystis* DHNA-CoA thioesterase knockout, and in fact are orthologous to recently characterized tomato methylketone synthases involved in the biosynthesis of 3-ketoacid volatiles (41). The fourth one -At1g68280, a paralog of At1g68260- corresponds to a pseudogene.

DHNA-CoA thioesterase activity occurs exclusively in peroxisomes, thus establishing definitive evidence for a split of the phylloquinone biosynthetic pathway between this organelle and plastids. This arrangement is apparently not specific to *Arabidopsis*, for DHNA-CoA thioesterase activity is not detectable either in plastids isolated from pea seedlings –one of the best sources for obtaining intact and highly pure chloroplasts (42)- while it is readily measured in the initial whole extract (Table 3-4). In addition, orthologs of AtDHNAT in monocots, gymnosperms, mosses and Lycopodiophytes all display canonical peroxisomal targeting signals type 1 (Fig. 3-7). Should it be confirmed that peroxisomes also bear the preceding DHNA-CoA synthase activity, the OSB and/or OSB-CoA precursors would have to exit plastids and enter peroxisomes to be converted into DHNA (Fig. 3-9). The present study implies that DHNA has then to accomplish the opposite path for the final phytylation and methylation steps. Why plants would have OSB-CoA ligases dual targeted to plastids and peroxisomes as recent experimental evidence indicates (13, 20) is a mystery. One hypothesis is that, since anions of hydrophobic weak acids flow in and out of chloroplasts, and consequently other organelles, in response to changes in stromal pH (43), a pool of OSB could passively diffuse into peroxisomes. The peroxisomal form of OSB-CoA ligase

could therefore be seen in this case as a salvaging enzyme. It is doubtful, on the other hand, that the hydrophilic OSB-CoA thioester, and DHNA, whose carboxyl group is >99% unprotonated at physiological pH (calculated pKas are 2.4 and 4.8 for the quinol and quinone forms, respectively), distribute themselves spontaneously between organelles. Specifically for DHNA, an *a priori* reason against passive diffusion is the redox active nature of its naphthoquinone ring. In other words, by freely shuttling through biological membranes, DHNA would act as an uncoupler and dissipate electrochemical gradients. A tacit conclusion is that yet-to-be identified transport steps between plastids and peroxisomes are likely involved in the biosynthesis of phylloquinone.

Table 3-4. DHNA-CoA thioesterase activity in pea seedling crude extracts and percoll-purified chloroplasts. Assays contained 18-90 µg protein and 6 µM DHNA-CoA. Data are means of three replicates ± SE.

	nmol hr^{-1} mg^{-1} protein
Crude extract	6.1 ± 0.85
Chloroplasts	< 0.5

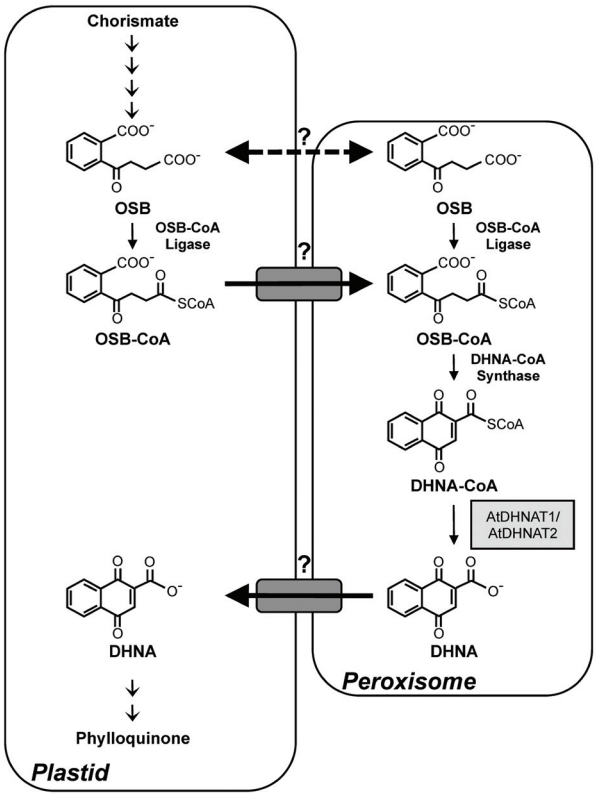


Fig. 3-9. Scheme of the probable trafficking of phylloquinone biosynthetic intermediates between plastids and peroxisomes. The dashed double-arrow indicates the hypothetical diffusion of OSB in and out of organelles. Solid arrows on grey boxes indicates predicted transport steps.

Analyses of cognate T-DNA knockout lines show that both AtDHNAT1 and AtDHNAT2 participate in phylloquinone biosynthesis in *Arabidopsis*, and that of the two, AtDHNAT1 is the prevailing contributor. Extracts of the double knockout *atdhnat1/atdhnat2*, however, still bear about 40% of the DHNA-CoA thioesterase activity of the wild-type, and depending on the growth conditions, 65% to 30% of its phylloquinone content. These results indicate that at least one more enzyme bears DHNA-CoA thioesterase activity in *Arabidopsis*. Figure 3-2 points to a candidate for such an additional enzyme –the *At5g48370* gene product– that partially complements the *Synechocystis slr0204* knockout. Remarkably, there is experimental evidence that the At5g48370 protein, like AtDHNAT1 and AtDHNAT2, is targeted to peroxisomes (44). One should also note that even in the *Synechocystis slr0204* knockout that is devoid of detectable DHNA-CoA thioesterase activity, the chemical hydrolysis of DHNA-CoA sustains a basal formation of phylloquinone representing 10-20% of the wild-type pool (22).

Besides invalidating their classification as PaaI enzymes, a structure-adjusted phylogenetic reconstruction established plant DHNA-CoA thioesterases as new functional members of the TE11/4HBT-II-type subfamily of hotdog-fold CoA thioesterases. That plant and cyanobacterial DHNA-CoA thioesterases display similar substrate specificity while having radically different catalytic sites exemplifies that, in the hotdog-fold superfamily, functions cannot be assigned on the basis of a mere comparison of primary sequences if the organisms are phylogenetically too distant. Most importantly, unlike their rhodophytes and euglyphidae counterparts, plant DHNA-CoA thioesterases are not of cyanobacterial ancestry. Instead, they are monophyletic with *Lactobacillales* orthologs that are encoded in clusters of vitamin K biosynthetic genes. It therefore seems very probable that the nuclearencoded plant DHNA-CoA thioesterase originates from an event of horizontal gene transfer with a menaquinone-synthesizing bacterium of the *Lactobacillales* order. There is actually a similar precedent for this in plant phylloquinone biosynthesis with OSB-CoA ligase, which appears to have been directly acquired from a δ-proteobacterium (45). We therefore propose that the early plastid-containing common ancestor of rhodophytes, green algae, and land plants, bore a cyanobacterial DHNA-CoA thioesterase of the TE12/Slr0204-type. This gene would have been lost or have significantly diverged after the

evolutionary split with red algae as it is no longer detected in green algae and land plants. Green algae that do not appear to contain orthologs of the TE11/4HBT-II-type enzyme presumably possess a third type of DHNA-CoA thioesterase.

Footnotes

This work was made possible in part by National Science Foundation Grant MCB-0918258 (to G.J.B.), support from the Center for Plant Science Innovation, and a graduate research assistantship from the Department of Agronomy and Horticulture at UNL (to J.R.W.). We thank Dr. Nicola Harrison-Lowe for her assistance with the preparation of peroxisomes, and Dr. Guodong Ren for his assistance with the making of the *Arabidopsis* double knockout.

References

1. Brettel, A., Sétif, P., and Mathis, P. (1987) *FEBS Lett.* 203, 220-224
2. Sigfridsson, K., Hansson, O., and Brzezinski, P. (1995) *Proc. Natl. Acad. Sci. USA* 92, 3458–3462
3. Singh, A. K., Bhattacharyya-Pakrasi, M., and Pakrasi, H. B. (2008) *J. Biol. Chem.* 283, 5762-15770
4. Furt, F., Oostende, C. V., Widhalm, J. R., Dale, M. A., Wertz, J., and Basset, G. J. C. (2010) *Plant J.* 64, 38-46
5. Yoshida, E., Nakamura, A., and Watanabe, T. (2003) *Anal. Sci.* 19, 1001-1005
6. Ikeda, Y., Komura, M., Watanabe, M., Minami, C., Koike, H., Itoh, S., Kashino, Y., and Satoh, K. (2008) *BBA-Bioenergetics* 1777, 351-361
7. Collins, M. D. and Jones, D. (1981) *Microbiol. Rev.* 45, 316-354
8. Booth, S. L. (2009) *Annu. Rev. Nutr.,* 29, 89-110
9. Booth, S. L., and Suttie, J. W. (1998) *J. Nutr.* 128, 785-788
10. Gross, J., Cho, W. K., Lezhneva, L., Falk, J., Krupinska, K., Shinozaki, K., Seki, M., Herrmann, R.G., and Meurer, J. (2006) *J. Biol. Chem.* 281, 17189-17196
11. Shimada, H., Ohno, R., Shibata, M., Ikegami, I., Onai, K., Ohto, M. A., and Takamiya, K. (2005) *Plant J.* 41, 627-637
12. Lohmann, A., Schottler, M.A., Brehelin, C., Kessler, K., Bock, R., Cahoon, E. D., and Dormann, P. (2006) *J. Biol. Chem.* 281, 40461-40472
13. Kim, H. U., van Oostende, C., Basset, G. J. and Browse, J. (2008) *Plant J.* 54, 272-283
14. Schultz, G., Ellerbrock, B., and Soll, J. (1981) *Eur. J. Biochem.* 117, 329-332
15. Kaiping, S., Soll, J., and Schultz, G. (1984). *Phytochemistry* 23, 89-91
16. Gaudillière, J.-P., d'Harlingue, A., Camara, B., and Monéger, R. (1984) *Plant Cell Rep.* 3, 240-242
17. Strawn, M. A., Marr, S. K., Inoue, K., Inada, N., Zubieta, C., and Wildermuth, M. C. (2007) *J. Biol. Chem.* 282, 5919-5933.
18. Garcion, C., Lohmann, A., Lamodière, E., Catinot, J., Buchala, A., Doermann, P., and Métraux J. P. (2008) *Plant Physiol.* 147, 1279-1287
19. Reumann, S., Babujee, L., Ma, C., Wienkoop, S., Siemsen, T., Antonicelli, G. E., Rasche, N., Lü der, F., Weckwerth, W., and Jahn, O. (2007) *Plant Cell* 19, 3170-3193
20. Babujee, L., Wurtz, V., Ma, C., Lueder, F., Soni, P., van Dorsselaer, A., and Reumann, S. (2010) *J. Exp. Bot.* 61, 1441-1453

21. Sakuragi, Y., and Bryant, D. A. (2006) *Photosystem I: The Light-driven Plastocyanin:Ferredoxin Oxidoreductase in Photosynthesis*, Springer, The Netherlands, Dordrecht
22. Widhalm, J. R., van Oostende, C., Furt, F., and Basset, G. J. (2009) *Proc. Natl. Acad. Sci. USA* 106, 5599-5603
23. Dillon, S.C., and Bateman, A. (2004) *BMC Bioinformatics,* 5, 109
24. Cantu, D. C., Chen, Y., and Reilly, P. J. (2010) *Protein Sci.* 19, 1281-1295
25. Benning, M. M., Wesenberg, G., Liu, R., Taylor, K. L., Dunaway-Mariano, D., and Holden, H. M. (1998) *J. Biol. Chem.* 273, 33572-33579
26. Thoden, J. B., Zhuang, Z., Dunaway-Mariano, D., and Holden, H. M. (2003) *J. Biol. Chem.* 278, 43709-43716
27. Jiang, M., Chen, X., Guo, Z. F., Cao, Y., Chen, M., and Guo, Z. (2008) *Biochemistry* 47, 3426-3434
28. Weigel D., and Glazebrook, J. (2002) *Arabidopsis: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
29. Reumann, S., Quan, S., Aung, K., Yang, P., Manandhar-Shrestha, K., Holbrook, D., Linka, N., Switzenberg, R., Wilkerson, C. G., Weber, A. P., Olsen, L. J. and Hu, J. (2009) *Plant Physiol.* 150, 125-143
30. Oostende, C., Widhalm, J. R., and Basset, G. J. (2008) *Phytochemistry* 69, 2457-2462
31. Sattler, S. E., Cahoon, E. B., Coughlan, S. J., and DellaPenna, D. (2003) *Plant Physiol.* 132, 2184-2195
32. Williams, J. G. K. (1988) *Methods Enzymol.* 167, 766-778
33. Karimi, M., Inze, D., and Depicker, A. (2002) *Trends Plant Sci.* 7, 193-195
34. Forner, J., and Binder, S. (2007) *BMC Plant Biol.* 7, 28
35. Hajdukiewicz, P., Svab, Z., and Maliga, P. (1994) *Plant Mol. Biol.* 25, 989-994
36. Finn R. D., Mistry, J., Tate, J., Coggill, P., Heger, A., Pollington, J. E., Gavin, O. L., Gunasekaran P., Ceric, G., Forslund, K., Holm, L., Sonnhammer, E.L., Eddy, S. R., and Bateman, A. (2010) *Nucleic Acids Res.* 38, D211-D222
37. Reumann, S. (2004) *Plant Physiol.* 135, 783-800
38. Reumann, S., Ma, C., Lemke, S., and Babujee, L. (2004) *Plant Physiol.* 136, 2587-2608
39. Sessions, A., Burke, E., Presting, G., Aux, G., McElver, J., Patton, D., Dietrich, B., Ho, P., Bacwaden, J., Ko, C., Clarke, J. D., Cotton, D., Bullis, D., Snell, J., Miguel, T., Hutchison, D., Kimmerly, B., Mitzel, T., Katagiri, F., Glazebrook, J., Law, M. and Goff, S. A. (2002) *Plant Cell,* 14, 2985-2994
40. Cantu, D. C., Chen, Y., Lemons, M. L., and Reilly, P. J. (2010) *Nucleic Acids Res.* 39, D342-D346
41. Yu, G., Nguyen, T. T. H., Guo, Y., Schauvinhold, I., Auldridge, M. E., Bhuiyan, N., Ben-Israel, I., Iijima, Y., Fridman, E., Noel, J. P., and Pichersky, E. (2010) *Plant Physiol.,* 154, 67-77
42. Cline, K. (1986) *J. Biol. Chem.* 261, 14804-14810
43. Cowan, I. R., Raven, J. A., Hartung, W., and Farquhar, G. D. (1982) *Aust. J. Plant Physiol.* 9, 489-498
44. Fukao, Y., Hayashi, M., and Nishimura, M. (2002) *Plant Cell Physiol.* 43, 689-696
45. Gross, J., Meurer, J., and Bhattacharya, D. (2008) Evidence of a chimeric genome in the cyanobacterial ancestor of plastids. *BMC Evol. Biol.,* 8, 117

CHAPTER 4

Re-Routing the Peroxisomal Branch of Phylloquinone Biosynthesis Back to Plastids

Introduction

Phylloquinone (2-methyl-3-phytyl-1,4-naphthoquinone; vitamin K_1) is synthesized by plants and various cyanobacteria (1) to serve as an electron carrier in photosystem I (2). A structurally-similar molecule, menaquinone (2-methyl-3-[all-trans-poly-isopentenyl]-1,4-naphthoquinone; vitamin K_2) is produced by some bacteria (3) and archaea as an electron transporter in anaerobic respiration (4) or purple-bacterial photosynthesis (5). Certain types of cyanobacteria (1), red algae (6), and diatoms (7) also synthesize menaquinone, but use it instead to fulfill the role of phylloquinone in photosystem I.

The formation of the naphthoquinone ring in menaquinone-synthesizing prokaryotes occurs through one of two routes starting from chorismate: an alternative pathway recently identified in some types of bacteria and archaea going through futalosine, a conjugate of chorismate and inosine (9) –or– the classically-described pathway, which relies on cyclization of the CoA-activated succinyl side chain of *o*-succinylbenzoate (OSB). All photosynthetic eukaryotes and cyanobacteria possess the latter route. This pathway has been elucidated primarily through identification of the *men* genes in *Escherichia coli* (3, 10, 11), which were subsequently used for homology-based searches to identify many of the corresponding orthologs used for phylloquinone biosynthesis in plants (12, 13, Chapter 3) and cyanobacteria (14-16).

The prevailing theory has long been that plants obtained their phylloquinone biosynthetic genes through engulfment of the ancient cyanobacterium, which later became the plastid. These genes were then transferred to the nucleus and the cognate enzymes targeted back to the plastid. Phylogenetic reconstructions now suggest that after the split with red algae, some plant phylloquinone biosynthetic genes were later acquired from menaquinone-producing bacteria through events of horizontal gene transfer (17, Chapter 3). Moreover, it appears that an all-plastidial model for phylloquinone biosynthesis in plants may be incorrect. New evidence

indicates that some of the genes involved in the production of the naphthoquinone ring actually encode enzymes localized to the peroxisome (18, 19, Chapter 3). Transient expression experiments show that OSB-CoA ligase (orthologous to *E. coli* MenE) is targeted to both plastids (13) and peroxisomes (19). Meanwhile, the 1,4-dihydroxy-2-naphthoyl-CoA (DHNA-CoA) synthase (MenB; 18, 19) and thioesterases (AtDHNAT1/2; Chapter 3) appear to be solely peroxisomal enzymes. Thus, the current model for plant phylloquinone biosynthesis reflects a split between plastids and peroxisomes. This implies, regardless of the true localization of OSB-CoA ligase, that there is movement of intermediates between plastids and peroxisomes. To assess the importance of the (unknown) transport steps in the pathway, we have re-routed the peroxisomal branch of phylloquinone biosynthesis by over-expressing bacterial versions of OSB-CoA ligase, DHNA-CoA synthase and thioesterase in the plastids of *Camelina sativa*. Here we report the findings of our metabolic engineering strategy.

Materials and Methods

Chemicals and Reagents: Synthesis of DHNA-CoA was performed as described in Chapters 2 and 3 using the modified protocol from (11). DHNA and menaquinone-4 were from Sigma, phylloquinone was from MP biomedicals (Solon, OH), and all other chemicals were from Fisher Scientific unless otherwise noted.

Plasmid Constructions: The cognate genes corresponding to *E.* coli MenE (NP_416763) and MenB (NP_416765), and *Synechocystis* Slr0204 (NP_442358), were synthesized by GenScript Corporation (Piscataway, NJ, U.S.). Each gene was codon optimized for expression in plants and designed to encode the N-terminal chloroplast targeting peptide of the *Coriandrum sativum* type II acyl-ACP desaturase (20). The *menB* and *slr0204* genes were designed to contain 5'-*EcoRI* and 3'-*BamHI* sites for directional cloning into the primary cloning vector pART7-ASC (21; Fig. 4-1A). The two resulting plasmids contained open reading frames of the *menB* or *slr0204* genes flanked by the cauliflower mosaic virus (CaMV) 35S promoter and octopine synthase terminator (Fig. 4-1A). These plant expression cassettes were then dropped from the pART7-

ASC vectors using *AscI* and subcloned into the *MluI* and *AscI* sites of the binary vector pBinRed2 (Fig. 4-1B), to generate the plasmid pBinRed2:BS. The menE gene was subcloned into pB2GW7 (22) using Gateway™ technology to create pB2GW7:E.

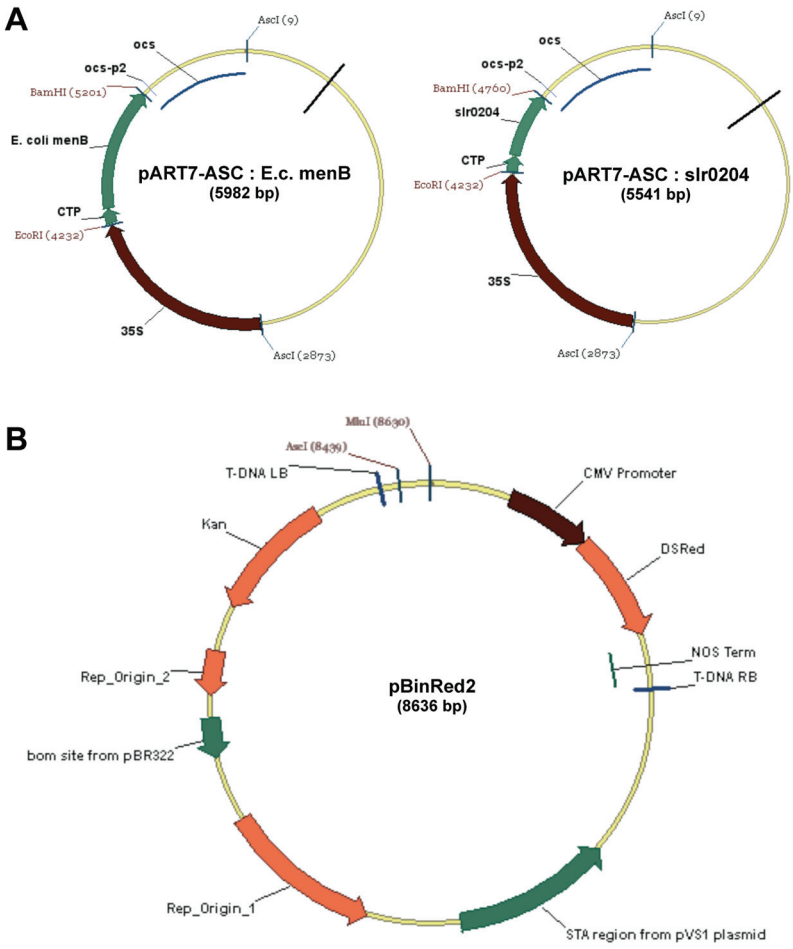


Fig. 4-1. Plasmids used to generate transgenic *Camelina* over-expressing *E. coli* menB and slr0204. *A,* pART7-ASC primary cloning vectors containing codon-optimized *E. coli* menB and slr0204 genes with 5' chloroplast targeting peptide (CTP) coding sequences in between the CaMV 35S promoter and octopine synthase (osc) terminator. *B*, Binary vector containing *MluI* and *AscI* sites for insertion of *E. coli* menB and slr0204 plant expression cassettes from pART7-ASC.

Plant Material and Generation of Transgenics: *Camelina* plants were grown in the greenhouse (24°C, 16-h days) from seeds germinated on soil. Plants were transformed via a floral dip method previously described (23) using *Agrobacterium tumefaciens* C58C1 electroporated with the described plasmid constructions. For the first round of transgenics, wild-type *Camelina* plants

were transformed with pBinRed2:BS. A control population was also created by transforming with the empty pBinRed2 vector. The T_1 seeds harvested from transformed plants were screened for expression of the DsRed fluorescent protein by illumination with a green LED flashlight and visualization with a red lens (23). Positive pBinRed2:BS T_1 seeds were used to grow the T_2 generation for transformation with pB2GW7:E. The T_2 seeds were first selected for red fluorescence, and then screened for glufosinate resistance on soil by spraying a 0.003% solution on seedlings after the emergence of the first true leaves. Glufosinate application was repeated two more times at three day intervals. Plants surviving the treatment were transplanted to new soil and screened by semi-quantitative RT-PCR for over-expression of the *menE*, *menB*, and *slr0204* genes.

Semi-quantitative RT-PCR: Total RNA from *Camelina* leaves was extracted using the SV Total RNA Isolation System (Promega) according to the manufacturer's recommendations. PCR (27 cycles) was conducted on cDNA prepared from 500 ng of total RNA using the following primers: *menE*, 5'- GCTCACGCTGCAACATGG-3' (forward) and 5'-ACAGAGGATCTGTTAACGAG-3' (reverse); *menB* 5'-AGGAACGCTTTCAGACCAC-3' (forward) and 5'-GAGGCACAACAGTATTAACC-3' (reverse); *slr0204* 5'-GCTGATACAGATGGAGCAG-3' (forward) and 5'-GGCTGTGGGAGTGGTGC-3' (reverse); *Actin* internal control, 5'-CTAAGCTCTCAAGATCAAAGGC-3' (forward) and 5'-TTAACATTGCAAAGAGTTTCAAGG-3' (reverse).

Phylloquinone Analysis: Phylloquinone content was measured on 10-30 mg of leaf tissue of four to six-week old plants using the extraction method and HPLC-fluoremetry detection protocol described in Chapter 3.

Results

In order to determine the effect of moving the peroxisomal branch of phylloquinone biosynthesis to plastids, transgenic *Camelina* plants were constructed that over-express plastid-targeted

bacterial versions of the OSB-CoA ligase (*E. coli menE*), DHNA-CoA synthase (*E. coli menB*), and DHNA-CoA thioesterase (*Syncechocystis slr0204*) genes (Fig. 4-2). Since OSB-CoA ligase is predicted to be dual targeted in plants (13, 19), we generated two types of transgenics. The first lines (*BS*-lines) over-express only *menB* and *slr0204* in plastids. Here, there is a reliance on the dual-targeted plant OSB-CoA ligase to produce the OSB-CoA intermediate. For the second round of transgenics (*EBS*-lines), the *menE* gene was also over-expressed in plastids already containing *menB* and *slr0204*.

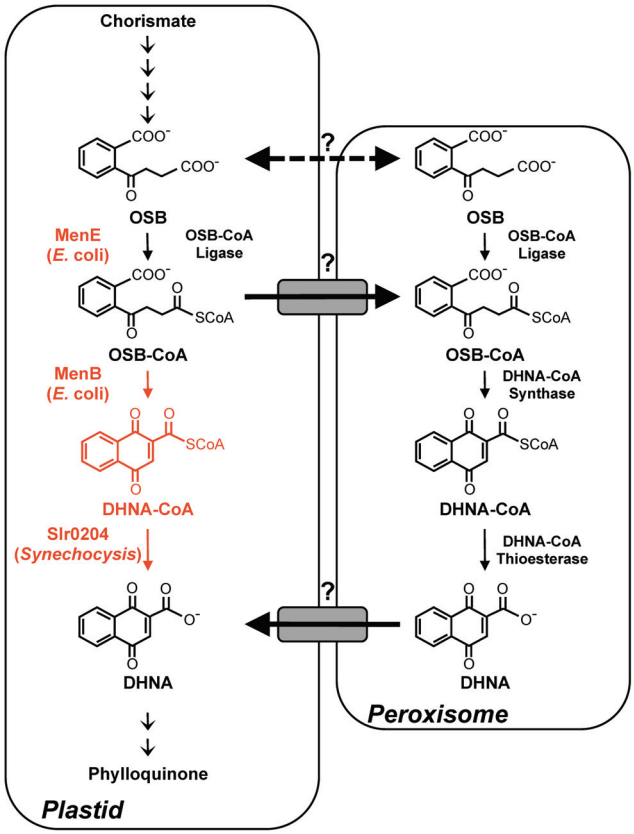


Fig. 4-2. Re-routing the peroxisomal branch of phylloquinone biosynthesis to plastids. Two transgenic approaches were taken. For the first, *menB* and *slr0204* were over-expressed in *Camelina* plastids. For the second, *menE*, *menB*, and *slr0204* were over-expressed in *Camelina* plastids. Shown in black are the assumed endogenous plant enzymes and intermediates. Shown in red are the predicted results of the engineering strategies.

Expression of *E. coli menB* and *Synechocystis slr0204* in Plastids Does Not Increase Phylloquinone Production: Wild-type *Camelina* plants were transformed with the construct pBinRed2 (empty vector; Fig. 4-1B) or pBinRed2:BS. Positive T_1 transformants (seeds selected for DsRed fluorescence) were sown on soil and screened by semi-quantitative RT-PCR to check for transgene expression. None of the analyzed plants coming from the empty-vector lines

showed expression of the *menB* and *slr0204* transgenes (Fig. 4-3A). On the other hand, all eight of the individual "*BS*-lines", those resulting from the transformation with *Agrobacteria* carrying pBinRed2:BS, showed expression of *menB* and *slr0204* (Fig. 4-3A). Next, we measured the phylloquinone content in leaves of six-week old plants coming from the control and *BS*-lines. On eight lines we found that control plants contained an average of 19.5 pmol phylloquinone mg^{-1} fresh weight (FW), with a range from 17 to 25.6 pmol phylloquinone mg^{-1} FW (Fig. 4-3B). Similarly, we found that on average the *BS*-lines examined contained 18.6 pmol phylloquinone mg^{-1} FW, ranging from 14.4 to 24.8 pmol phylloquinone mg^{-1} FW (Fig. 4-3B). Thus, there appears to be no increase in phylloquinone content gained from over-expressing *menB* and *slr0204* in *Camelina* plastids.

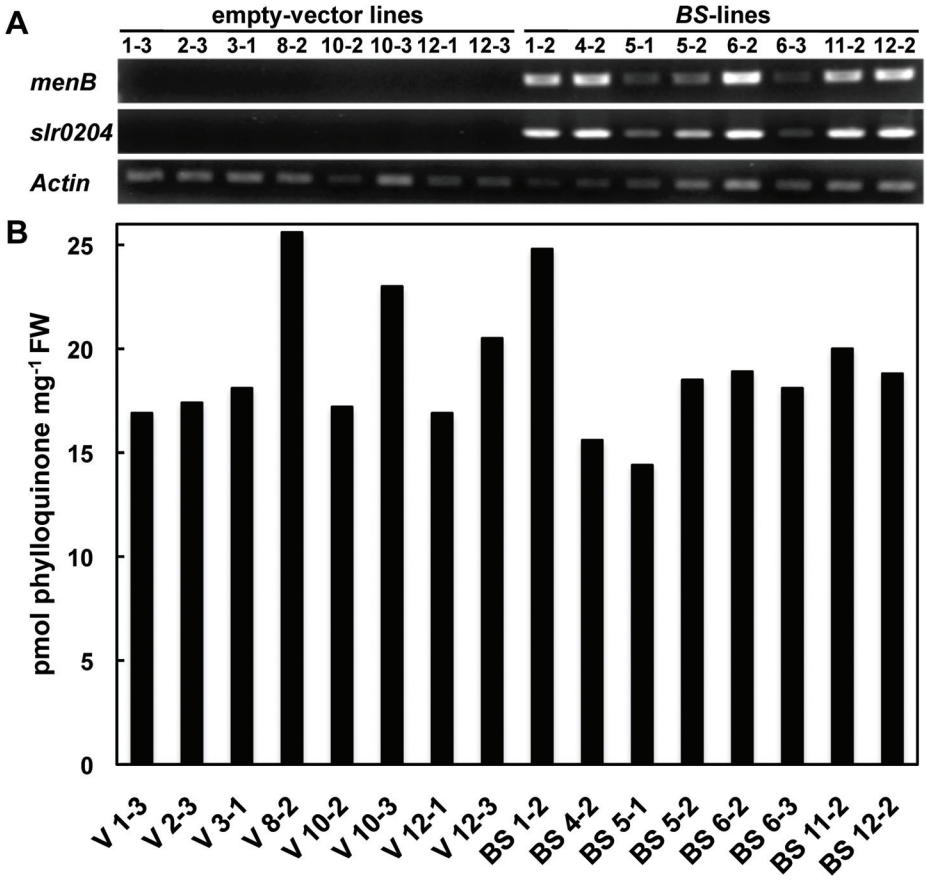


Fig. 4-3. Transgene expression and phylloquinone content of empty-vector and *BS*-lines. *A,* Semi-quantitiative RT-PCR analysis of *menB*, *slr0204* and *Actin* (internal control) expression in leaves of *Camelina* empty-vector (V) and *BS*-lines (see text for description of lines). *B,* Phylloquinone content in six-week old leaves of *Camelina* empty-vector and *BS*-lines. The expression analysis for each line in panel *A* corresponds to the phylloquinone content reported in panel *B*.

Expression of *E. coli menE* in Plastids of Transgenic *BS*-lines Results in a Limited Increase of Phylloquinone: Five *BS*-lines (1-2, 3-1, 4-2, 5-2, and 6-2) were transformed with the construct pB2GW7:E described in Materials and Methods. Positive T_1 transformants were selected for glufosinate resistance at the seedling stage and transplanted to fresh soil. Leaves from approximately four-week old plants, representing 42 independent *EBS*-lines and four empty-vector lines, were first screened for phylloquinone content (Fig. 4-4). Across four lines, empty-vector control plants contained an average of 11.6 pmol phylloquinone mg^{-1} FW, ranging from 9 to 15 pmol phylloquinone mg^{-1} FW. The 42 *EBS*-lines contained an average of 11.9 pmol phylloquinone mg^{-1} FW, ranging from 7.8 to 16.8 pmol phylloquinone mg^{-1} FW. Thus, as a whole, showing no significant increase in phylloquinone production above the control lines. However, three of the *EBS*-lines (4-2-3, 4-2-4, and 5-2-9) showed an approximately 28% increase in phylloquinone content (Table 4-1). Semi-quantitative RT-PCR showed that plants from these lines expressed the *menE*, *menB*, and *slr0204* transgenes, while the vector-alone counterparts did not (Fig. 4-5).

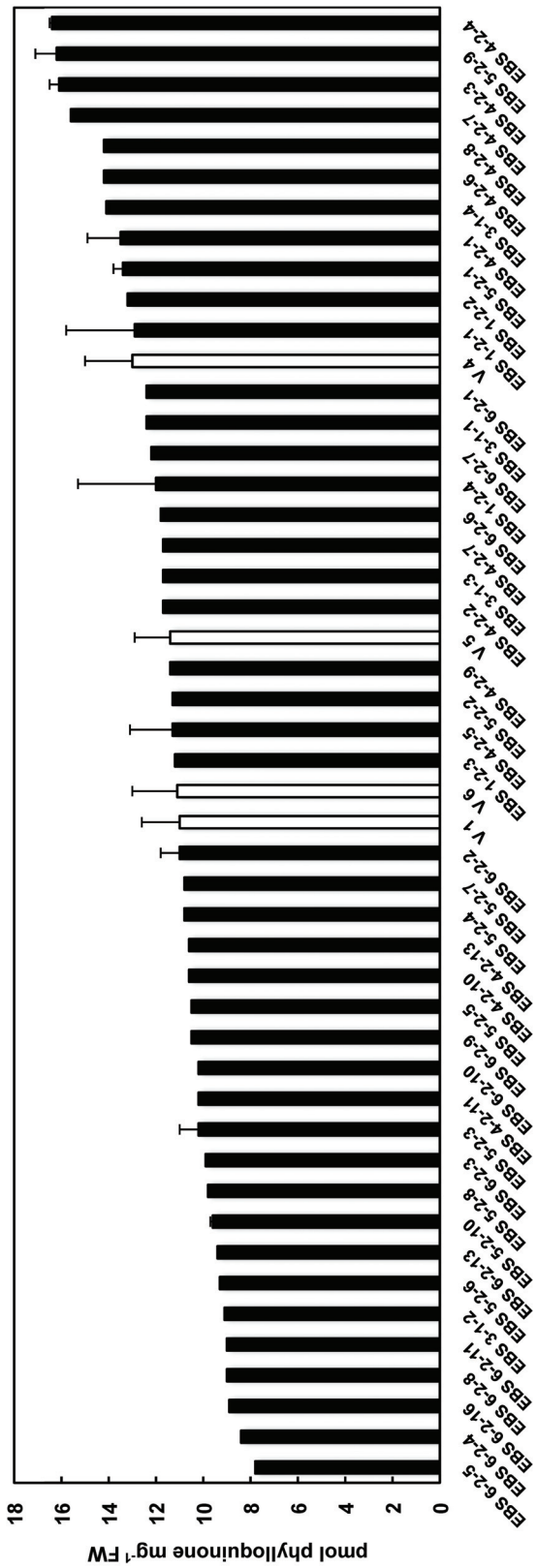


Fig. 4-4. Phylloquinone content of empty-vector and E*BS*-lines. Phylloquinone content in four-week old leaves of *Camelina* empty-vector (V, open bars) and *EBS*-lines (filled bars). See text for description of lines. Data with error bars are means ± SE of at least two replicates.

Table 4-1 Phylloquinone content of select empty-vector and E*BS*-lines. Phylloquinone content in four-week old leaves of *Camelina* empty-vector (V) and *EBS*-lines (see text for description of lines). Data are means ± SE of at least two replicates.

	pmol phylloquinone mg^{-1} FW	
Line	Average	Average across all lines
V 1	11.0 ± 1.6	11.6 ± 0.9
V 6	11.1 ± 1.9	
V 5	11.4 ± 1.5	
V 4	13.0 ± 2.0	
EBS 4-2-3	16.1 ± 0.4	16.2 ± 0.1
EBS 5-2-9	16.2 ± 0.9	
EBS 4-2-4	16.4 ± 0.1	

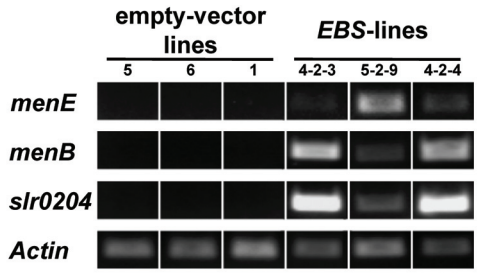


Fig. 4-5. Transgene expression of select empty-vector and *EBS*-lines. Semi-quantitiative RT-PCR analysis of *menE*, *menB*, *slr0204* and *Actin* (internal control) expression in leaves of *Camelina* empty-vector (V) and *EBS*-lines (see text for description of lines).

Discussion

It has recently been discovered that phylloquinone biosynthesis in land plants is split between plastids and peroxisomes (18, 19, Chapter 3). Inherent in this architecture is the need for intermediate trafficking between organelles. The hydrophilic nature of CoA conjugates, and the redox nature of the naphthoquinone ring (see Chapters 1 and 3 for further explanation), make it likely that regulated transport steps are required to shuttle the metabolites between plastids and peroxisomes. In this study we attempted to bypass these putative transport steps (Fig. 4-2) by over-expressing in plastids the bacterial versions of the peroxisomal phylloquinone biosynthetic enzymes. In our first engineering attempt we targeted *E. coli menB* (DHNA-CoA synthase) and *Synechocystis slr0204* (DHNA-CoA thioesterase) to plastids in *Camelina*. Across all lines we saw no significant increase in phylloquinone production compared to the empty-vector controls with this strategy (Fig. 4-3A).

In light of the recent demonstration that OSB-CoA ligase in plants may be dual-targeted to both plastids (13) and peroxisomes (19), we hypothesized that perhaps this reaction could be limiting our approach. A closer look at the subcellular localization strategies proposing plastidial (13) and peroxisomal (19) targeting of plant OSB-CoA ligase reveals biases in the experiments. By fusing GFP to the N- and C-terminal ends of the protein the putative chloroplast targeting peptide and peroxisomal targeting signal type I (PTS1), respectively, are masked. Thus, it is possible that OSB-CoA ligase could be preferably targeted to one of the organelles. To rule out that OSB-CoA ligase is creating a bottleneck in our transgenic strategy, we undertook a second round of engineering to over-express *E. coli menE* in *Camelina* plastids. In 39 of the 42 lines analyzed we saw no significant increase in phylloquinone content (Fig. 4-4). In three *EBS*-lines, however, we detected an increase in phylloquinone content roughly 28% higher than the empty-vector control lines (Fig. 4-4, Table 4-1). Although statistically significant, this moderate increase in phylloquinone content could simply be due to natural variation among plants. Especially since we only see it in 7% of the analyzed lines. It may be helpful to screen several more control lines to rule out natural variation. Moreover, phylloquinone appears to accumulate in leaves as they mature (unpublished data), leaving ample room for error; great care was taken to ensure systematic sampling in this study to avoid this effect. Leaf maturity does likely explain, however, why phylloquinone content measured in *BS*-lines and their controls (Fig. 4-3B) were higher than the *EBS*-lines and their controls (Fig. 4-4). Finally, despite the fact that semi-quantitative RT-PCR confirmed expression of the transgenes (Fig. 4-3A, Fig. 4-5), Western blot analyses should be performed on plastidial extracts to ensure targeting and adequate expression of the MenE, MenB, and Slr0204 proteins.

One collective feature of the peroxisomal steps of phylloquinone biosynthesis is that they catalyze the CoA reactions of the pathway. It is suggested that 88-100% of CoA in plastids is in the form of acetyl-CoA (precursor for lipid biosynthesis) with a variable pool of free CoA often falling below 0.1 μM (reviewed in 24). Additionally, there is a large flux of CoA that leaves the plastid. Through action of acyl-CoA synthetase, acyl CoAs are transported out of the plastid to go to the ER to be incorporated into glycerolipid biosynthesis (25). Since the reactions to synthesize CoA occur in

the mitochondria and cytosol (reviewed by 26), coupled with the hydrophilic nature of CoA, there are likely transport steps that limit a compensatory influx of the cofactor into plastids. It is thus reasonable to hypothesize that our failure to increase phylloquinone content by moving the peroxisomal branch to plastids is due to an insufficient pool of free CoA, assuming that all of the control experiments mentioned earlier check out. It is also interesting to note that over-expressing bacterial isochorismate synthase (ICS, the enzyme catalyzing the first reaction in the pathway) resulted in a 400% increase in phylloquinone content (27). This successful engineering attempt *a priori* relies on the peroxisomal pool of CoA. This would seem to indicate then that the transport steps we attempted to bypass by re-routing the peroxisomal branch are not as limiting as the pool of isochorismate or the availability of free CoA in plastids. If the latter true, it may have served as the evolutionary pressure to move the CoA reactions in phylloquinone biosynthesis from the plastid to the peroxisome. One way to test this is to use the same engineering strategies presented here in phylloquinone biosynthetic mutants corresponding to enzymes of the peroxisomal branch. It may also be necessary to over-express the bacterial ICS to ensure a sufficient pool of isochorismate is present. Only then can it be determined if the plastidial branch is sufficient to compensate for the peroxisomal portion of phylloquinone biosynthesis.

References

1. Sakuragi Y & Bryant DA (2006) in ed Golbeck JH (Springer, Dordrecht, The Netherlands), pp 205-222.
2. Fromme P & Grotjohann I (2006) in ed Golbeck JH (Springer, Dordrecht, The Netherlands), pp 47-69.
3. Meganathan R (2001) Biosynthesis of menaquinone (vitamin K2) and ubiquinone (coenzyme Q): A perspective on enzymatic mechanisms. *Vitam Horm* 61, 173-218.
4. Collins MD & Jones D (1981) Distribution of isoprenoid quinone structural types in bacteria and their taxonomic implication. *Microbiol Rev* 45, 316-354.
5. Schoepp-Cothenet B*, et al* (2009) Menaquinone as pool quinone in a purple bacterium. *Proc Natl Acad Sci U S A* 106, 8549-8554.
6. Yoshida E, Nakamura A & Watanabe T (2003) Reversed-phase HPLC determination of chlorophyll a' and naphthoquinones in photosystem I of red algae: Existence of two menaquinone-4 molecules in photosystem I of cyanidium caldarium. *Anal Sci* 19, 1001-1005.
7. Ikeda Y*, et al* (2008) Photosystem I complexes associated with fucoxanthin-chlorophyll-binding proteins from a marine centric diatom, *Chaetoceros gracilis. Biochim Biophys Acta* 1777, 351-361.

8. Rodriguez-Concepcion M & Boronat A (2002) Elucidation of the methylerythritol phosphate pathway for isoprenoid biosynthesis in bacteria and plastids. A metabolic milestone achieved through genomics. *Plant Physiol* 130, 1079-1089.
9. Hiratsuka T, *et al* (2008) An alternative menaquinone biosynthetic pathway operating in microorganisms. *Science* 321, 1670-1673.
10. Jiang M, *et al* (2007) Menaquinone biosynthesis in escherichia coli: Identification of 2-succinyl-5-enolpyruvyl-6-hydroxy-3-cyclohexene-1-carboxylate as a novel intermediate and re-evaluation of MenD activity. *Biochemistry* 46, 10979-10989.
11. Jiang M, *et al* (2008) Identification and characterization of (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of *Escherichia coli. Biochemistry* 47, 3426-3434.
12. Gross J, *et al* (2006) A plant locus essential for phylloquinone (vitamin K1) biosynthesis originated from a fusion of four eubacterial genes. *J Biol Chem* 281, 17189-17196.
13. Kim HU, van Oostende C, Basset GJ & Browse J (2008) The AAE14 gene encodes the arabidopsis o-succinylbenzoyl-CoA ligase that is essential for phylloquinone synthesis and photosystem-I function. *Plant J* 54, 272-283.
14. Johnson TW, *et al* (2000) Recruitment of a foreign quinone into the A(1) site of photosystem I. I. genetic and physiological characterization of phylloquinone biosynthetic pathway mutants in *Synechocystis* sp. pcc 6803. *J Biol Chem* 275, 8523-8530.
15. Wade Johnson T, *et al* (2003) The menD and menE homologs code for 2-succinyl-6-hydroxyl-2,4-cyclohexadiene-1-carboxylate synthase and O-succinylbenzoic acid-CoA synthase in the phylloquinone biosynthetic pathway of *Synechocystis* sp. PCC 6803. *Biochim Biophys Acta* 1557, 67-76.
16. Widhalm JR, van Oostende C, Furt F & Basset GJ (2009) A dedicated thioesterase of the hotdog-fold family is required for the biosynthesis of the naphthoquinone ring of vitamin K1. *Proc Natl Acad Sci U S A* 106, 5599-5603.
17. Gross J, Meurer J & Bhattacharya D (2008) Evidence of a chimeric genome in the cyanobacterial ancestor of plastids. *BMC Evol Biol* 8, 117.
18. Reumann S, *et al* (2007) Proteome analysis of arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *Plant Cell* 19, 3170-3193.
19. Babujee L, *et al* (2010) The proteome map of spinach leaf peroxisomes indicates partial compartmentalization of phylloquinone (vitamin K1) biosynthesis in plant peroxisomes. *J Exp Bot* 61, 1441-1453.
20. Cahoon EB, Shanklin J & Ohlrogge JB (1992) Expression of a coriander desaturase results in petroselinic acid production in transgenic tobacco. *Proc Natl Acad Sci U S A* 89, 11184-11188.
21. Gleave AP (1992) A versatile binary vector system with a T-DNA organisational structure conducive to efficient integration of cloned DNA into the plant genome. *Plant Mol Biol* 20, 1203-1207.
22. Karimi M, Inze D & Depicker A (2002) GATEWAY vectors for agrobacterium-mediated plant transformation. *Trends Plant Sci* 7, 193-195.
23. Lu C & Kang J (2008) Generation of transgenic plants of a potential oilseed crop *Camelina sativa* by agrobacterium-mediated transformation. *Plant Cell Rep* 27, 273-278.
24. Daae EB, *et al* (1999) Metabolic modeling as a tool for evaluating polyhydroxyalkanoate copolymer production in plants. *Metab Eng* 1, 243-254.
25. Bowsher, C., Steer, M., Tobin, A, Cornelius, D. P., Rayburn, M., & Mantauk, B. B. (2008). *Plant biochemistry*. New York, NY: Garland Science.
26. Roje S (2007) Vitamin B biosynthesis in plants. *Phytochemistry* 68, 1904.
27. Verberne MC, Sansuk K, Bol JF, Linthorst HJ & Verpoorte R (2007) Vitamin K1 accumulation in tobacco plants overexpressing bacterial genes involved in the biosynthesis of salicylic acid. *J Biotechnol* 128, 72-79.

CHAPTER 5

Conclusions

The identification of DHNA-CoA thioesterases in cyanobacteria and plants described here in effect completes the elucidation of the enzymes involved in phylloquinone biosynthesis—in plants the DHNA-CoA synthase (MenB) ortholog has not been experimentally demonstrated, but expression data and homology strongly point to the *At1g60550* gene product. Nevertheless, it still remains that the DHNA-CoA thioesterase for menaquinone production in *E. coli* is unknown. One clue may come from the work of Jiang *et al* (1) in their effort to show that MenH is the SEPHCHC synthase and not the DHNA-CoA thioesterase. This group used an *in vitro* assay to demonstrate that MenH cannot hydrolyze DHNA-CoA by comparison to a positive control enzyme called P15 (1). It turns out that P15 is actually an *E. coli* gene product (YbdB; PDB structure 1VH9) that we found to be homologous to the 4HBT-II-type of DHNA-CoA thioesterase found in plants and encoded in the menaquinone biosynthetic operons of several *Lactobacillales* species (Fig. 3-8B). Unfortunately, the *E. coli* knockout does not lack menaquinone (Widhalm *et al.* unpublished data). From this, we predict that in addition to P15, there may be another enzyme(s) moonlighting on the substrate *in vivo*, reopening *de facto* the search for DHNA-CoA thioesterase(s) in *E. coli*.

It has been proposed that prior to endosymbiosis, the cyanobacterial plastid ancestor captured a *men* cluster through horizontal gene transfer (Fig. 1-8), which is today preserved as a chimeric relic in the plastid genomes of the red algae Cyanidales (2). This model is intended to explain how the *men* orthologs in red algae and plants could have been obtained through plastid endosymbiosis, yet be more closely related to modern Chlorobi/γ-proteobacteria than cyanobacteria. Nonetheless, using comparative genomics to look at modern cyanobacteria, Cyanidiales and Chlorobi/γ-proteobacteria there is something that this evolutionary model cannot explain. The genomic arrangement of the *men* and DHNA-CoA thioesterase homologs of many present-day cyanobacteria and Cyanidiales displays striking similarities that are not conserved in the Chlorobi/γ-proteobacteria lineage (Fig. 5-1). Although this conservation might be purely coincidental, or driven by identical selective constraints (e.g. transcriptional regulation), it could

also point to an overlooked phylogenetic relatedness between the vitamin K biosynthetic genes of Cyanidiales and cyanobacteria.

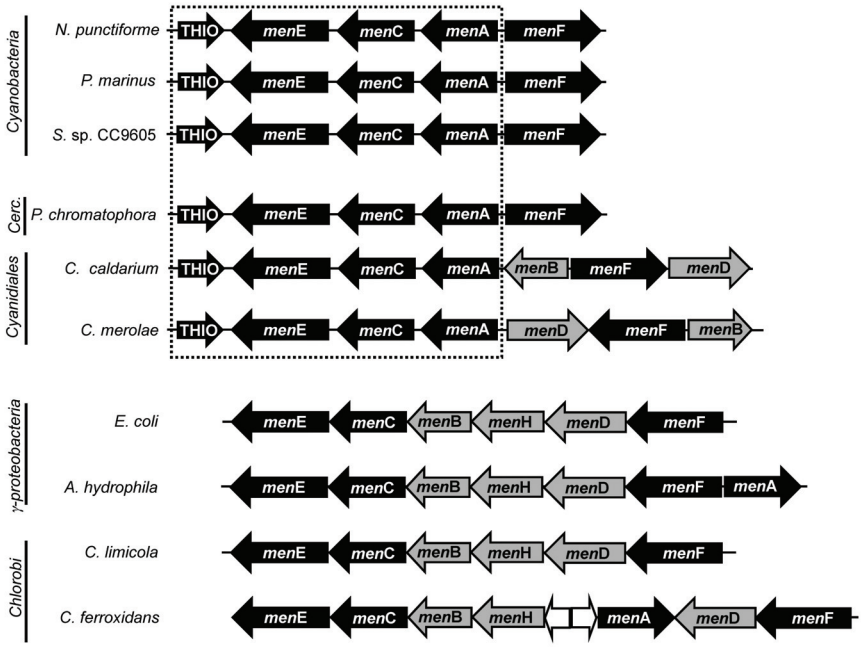


Fig. 5-1. Organization of the phylloquinone/menaquinone biosynthetic gene clusters in representative species of cyanobacteria/cercozoan, cyanidiales, γ-proteobacteria and chlorobi. The dashed frame highlights the conserved arrangement of the *men* and Slr0204-type DHNA-CoA thioesterase (THIO) homologs in cyanobacteria (*N. punctiforme* (*Nostoc punctiforme*); *P. marinus* (*Prochlorococcus marinus*); S. sp. CC9605 (*Synechoccoccus* sp. CC9605)), cercozoan (Cerc.) (*P. chromatophora* (*Paulinella chromatophora*)) and cyanidiales (*C. caldarium* (*Cyanidium caldarium*); *C. merolae* (*Cyanidioschyzon merolae*)). The gene cluster of the cercozoan species *P. chromatophora* is located in a plastid-like organelle called the chromatophore; the later is thought to originate from a recent endosymbiosis of a cyanobacterium of the *Prochlorococcus*/*Synechococcus* lineage. *A. hydrophila* (*Aeromonas hydrophila*); *C. limicola* (*Chlorobium limicola*); *C. ferroxidans* (*Chlorobium ferroxidans*).

The presence of an *slr0204* homolog in the *men* clusters of red algal plastid genomes (Fig. 5-1) suggests that the cyanobacterial-type of DHNA-CoA thioesterase existed in the ancestral plastid. Since we do not retrieve Slr0204 homologs in plants today, we can infer that it was lost sometime after the split between red and green algae. This event was likely just preceded by the horizontal gene transfer of the *Lactobacillales* 4HBT-II-type of DHNA-CoA thioesterase (Fig. 5-2). Note that this seems to be the case for plants—we still have not found either type of DHNA-CoA thioesterase in green algae (Fig. 5-2).

In Chapter 2 we proposed that there are four plastidial Slr0204-like proteins in plants that can be detected using PSI-BLAST searches. In actuality the *E*-values (e^{-4}) retrieved for these

candidates are fairly insignificant, and as we showed in Chapter 3, none of them could complement the cyanobacterial DHNA-CoA thioesterase mutant. Moreover, these plastidial enzymes have recently been shown to be methylketone synthases (3), which catalyze a reaction very similar to that of DHNA-CoA thioesterase. Based on some of the primary sequence resemblances (Fig. 2-6), and the similar chemistry used in the methylketone synthase mechanism, I am tempted to speculate that these are the "lost" Slr0204-type of DHNA-CoA thioesterases in plants. That is, after the acquisition of the 4HBT-II-type of DHNA-CoA thioesterase from *Lactobacillales*, the pressure to maintain the plastidial Slr0204-like version was "lost" allowing it to evolve a methylketone synthase function.

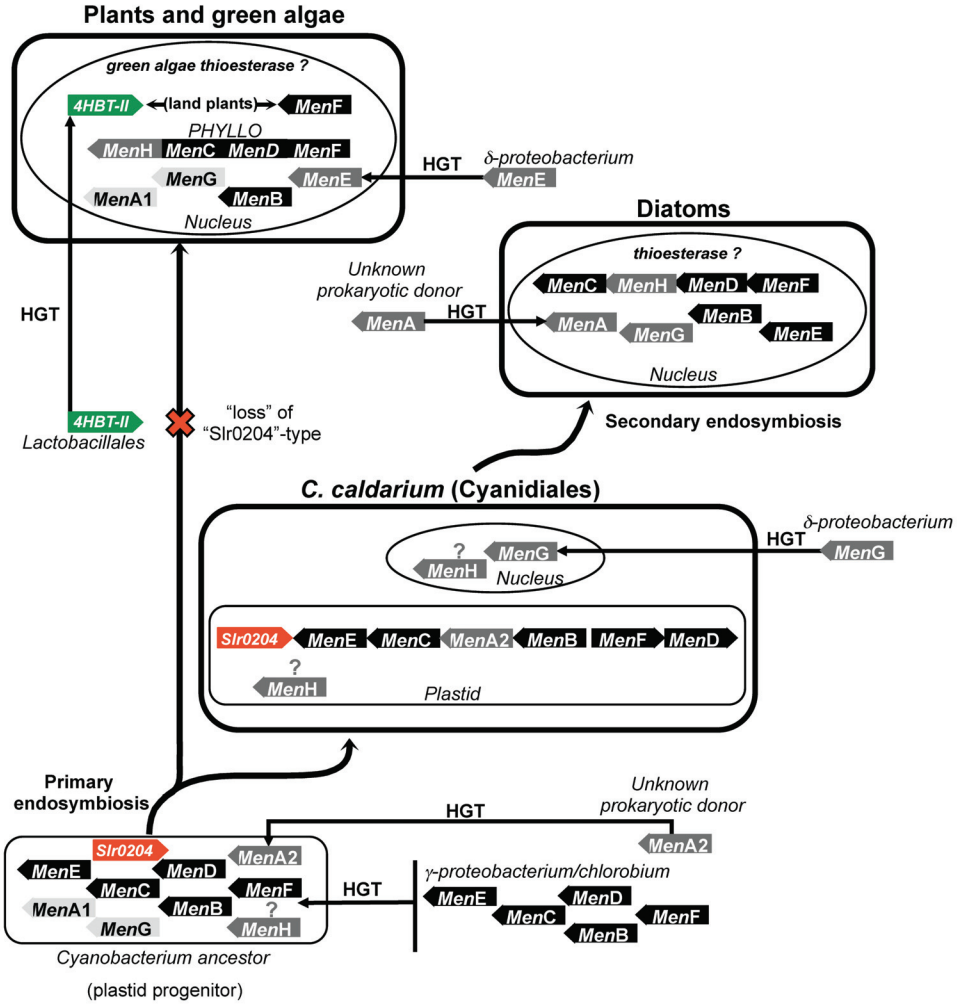


Fig. 5-2. Proposed evolution of DHNA-CoA thioesterase genes in photosynthetic eukaryotes. Appended to the model proposed by Gross et al (2008) shown in Figure 1-6. Modern red algae (Cyanidales) contain the cyanobacterial-type Slr0204 DHNA-CoA thioesterase (in red) within *men* clusters in the plastid genome resembling present-day cyanobacteria (Fig. 5-1). Land plants do not contain Slr0204 homologs, so we propose that they were "lost" around the time the *Lactobacillales* 4HBT-II-type of DHNA-CoA thioesterase (in green) was acquired. Note that green algae and diatoms contain neither Slr0204 nor 4HBT-II homologs, meaning they likely use another type of thioesterase. HGT, Horizontal Gene Transfer. See Chapter 1 and Figure 1-6 for further details regarding *men* gene evolution.

In our study we show that plant DHNA-CoA thioesterase is peroxisomal using both GFP-reporter and biochemical evidence. To be fair, though, we are not the first to dispute the all-plastidial localization of phylloquinone biosynthesis in plants. Reports of peroxisomal targeting signals hidden in the N-terminal end of putative plant MenB orthologs, and dual plastidial/peroxisomal-targeting signals in AAE14 (MenE), have been confirmed by GFP-reporter experiments in the last few years (4, 5). Coupled with our identification of DHNA-CoA thioesterase, we now have enough evidence to propose a model depicting a split in the plant phylloquinone biosynthetic pathway between plastids and peroxisomes (Fig. 3-9).

The constraints (or lack there of) favoring a compartmental split in plant phylloquinone biosynthesis are still under investigation. We can assume that this split does not occur in all photosynthetic eukaryotes because most of the *men* genes (including MenE, MenB, and DHNA-CoA thioesterase) in red algae are still encoded in the plastid genome. In chlorophytic green algae, such as *Chlamydomonas*, the MenB and MenE orthologs do not contain peroxisomal targeting signals (5), likely pointing to an all-plastidial pathway. Interestingly, in *Physcomitrella patens* (moss), the MenB ortholog also lacks a peroxisomal targeting signal (5), while the DHNA-CoA thioesterase ortholog bears a PTS1 (Chapter 3). Peroxisomal targeting peptides can be detected, however, in the MenE, MenB, and DHNA-CoA thioesterase orthologs in *Selaginella moellendorffii* ("spikemoss"). This seems to suggest that the compartmental split of phylloquinone biosynthesis between plastids and peroxisomes may be apomorphic to vascular plants.

In Chapter 4 we described an engineering attempt to move the peroxisomal branch of phylloquinone biosynthesis to the plastids. The split of phylloquinone biosynthesis between plastids and peroxisomes implies that there are transport steps necessary to move intermediates between the organelles. Though there are some crucial control experiments to be done to ensure

proper expression of the bacterial-versions of the peroxisomal enzymes in plastids, our metabolic engineering attempt (Chapter 4) to bypass these putative transport steps failed to consistently increase phylloquinone production. Interestingly, a previous strategy over-expressing the bacterial isochorismate synthase resulted in a 400% increase in phylloquinone production (6). This indicates that in our strategy the constraints of the cognate transport steps between plastids and peroxisomes are not as limiting as the pool of isochorismate. In our next round of engineering we plan to also introduced bacterial ICS to direct a larger flux of chorismate to phylloquinone biosynthesis. Moreover, an important observation that we made while considering a hypothesis for the organellar split in phylloquinone biosynthesis, is that the peroxisomal steps in the pathway are also those involved in the CoA reactions. We speculate that the plastidial pool of CoA could also be a limiting factor in our strategy. It is possible that the evolutionary pressure for moving the enzymes in question to peroxisomes was to force phylloquinone biosynthesis to tap into a non-plastidial pool of CoA. Since most CoA is tied up in the form of acetyl-CoA in plastids (7), presumably for lipid biosynthesis, there may have been limited quantities to fulfill the requirements of both vital pathways. To test this, we will perform our engineering strategy in phylloquinone biosynthetic mutants corresponding to enzymes of the peroxisomal branch. We predict that If CoA is limiting in plastid that there are likely other cyanobacterial-derived metabolic pathways with CoA reactions that have been moved from the plastid to a peroxisomal branch. Already, recent studies in petunia suggest that the CoA-dependent formation of benzoate (analogous to the β-oxidation of fatty acids) for flower volatile biosynthesis utilizes a peroxisomal branch (8). This pathway starts from the plastid-derived precursor phenylalanine, and may have evolved from an acquired cyanobacterial route to synthesize benzenoid compounds. Thus, as higher plants evolved the ability to produce more and more secondary metabolites, many of them proceeding through CoA intermediates, it may have become advantageous to centralize this chemistry in the peroxisome.

Final Thoughts

The exploration of DHNA-CoA thioesterase has resulted in many exciting discoveries, but more importantly has taught me about developing and executing a research project. When I started in the lab we had the notion that the α/β-hydrolase module of PHYLLO (the MenH domain) was the DHNA-CoA thioesterase. After this was disproven, we were left scratching our heads and back at the beginning. This turned out to be a defining moment in my education, because soon after I got a true research experience from the ground up on how to go about identifying and characterizing an unknown step in a metabolic pathway. Overall, the project developed quite nicely, and we went on to make some very exciting and unexpected findings. In addition to discovering the first DHNA-CoA thioesterase, we have determined that plants today no longer contain the cyanobacterial ortholog that was *a priori* inherited from the plastid endosymbiotic event. Instead, plants recently acquired a highly-diverged ortholog from a menaquinone-producing bacteria. Astonishingly, plants have also moved an entire branch of phylloquinone biosynthesis from the plastid to the peroxisome. In light of these discoveries we are faced with even more questions. Why have certain enzymes been replaced through horizontal gene transfers, and why is an entire branch of the pathway now in the peroxisome? Is one of these the result of the other? What are the factors involved in regulating the cross-talk between the peroxisomal and plastidial branches of the pathway? Does phylloquinone biosynthesis intersect with an unsuspected pathway?

References

1. Jiang M*, et al* (2008) Identification and characterization of (1R,6R)-2-succinyl-6-hydroxy-2,4-cyclohexadiene-1-carboxylate synthase in the menaquinone biosynthesis of escherichia coli*. Biochemistry* 47, 3426-3434.
2. Gross J, Meurer J & Bhattacharya D (2008) Evidence of a chimeric genome in the cyanobacterial ancestor of plastids*. BMC Evol Biol* 8, 117.
3. Yu G*, et al* (2010) Enzymatic functions of wild tomato methylketone synthases 1 and 2*. Plant Physiol* 154, 67-77.
4. Reumann S*, et al* (2007) Proteome analysis of arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms*. Plant Cell* 19, 3170-3193.
5. Babujee L*, et al* (2010) The proteome map of spinach leaf peroxisomes indicates partial compartmentalization of phylloquinone (vitamin K1) biosynthesis in plant peroxisomes*. J Exp Bot* 61, 1441-1453.
6. Verberne MC, Sansuk K, Bol JF, Linthorst HJ & Verpoorte R (2007) Vitamin K1 accumulation in tobacco plants overexpressing bacterial genes involved in the biosynthesis of salicylic acid*. J Biotechnol* 128, 72-79.
7. Daae EB*, et al* (1999) Metabolic modeling as a tool for evaluating polyhydroxyalkanoate copolymer production in plants*. Metab Eng* 1, 243-254.

8. Van Moerkercke A, Schauvinhold I, Pichersky E, Haring MA & Schuurink RC (2009) A plant thiolase involved in benzoic acid biosynthesis and volatile benzenoid production. *Plant J* 60, 292-302.